



Koivisto Aleksi

Yläkoulujen päättökeskiarvojen yhteys niiden lähialueiden sosioekonomisiin tekijöihin

Oulussa

Pro gradu -tutkielma
KASVATUSTIETEIDEN TIEDEKUNTA
Luokanopettajakoulutus
2020

Oulun yliopisto

Kasvatustieteiden tiedekunta

Yläkoulujen päättökeskiarvojen yhteys niiden lähialueiden sosioekonomisiin tekijöihin

Oulussa (Aleksi Koivisto)

Pro gradu -tutkielma, 114 sivua

Marraskuu 2020

Tässä pro gradu -tutkielmassa tarkastellaan Oulun kaupungin yläkoulujen oppimistuloksien yhteyksiä koulujen sijaintialueiden sosioekonomisiin tekijöihin. Tutkimuksessa selvitettiin Oulun yläkoulujen päättökeskiarvojen yhteyttä koulujen lähialueiden tulo- ja koulutustasoihin. Tutkimusaihe on ollut ajankohtainen esimerkiksi PISA-tulosten osoittaman koulujen eriytymiskehityksen vuoksi. Aiemmat tutkimukset ovat osoittaneet, että yhteyksiä koulujen lähialueiden ja oppimistulosten väliltä löytyy, mutta tutkimus on Suomessa keskittynyt pääosin pääkaupunkiseudulle. Tässä pro gradu -tutkielmassa pyrittiin laajentamaan aihetta käsittelevää tutkimusta pääkaupunkiseudun ulkopuolelle ja Pohjois-Suomeen.

Tutkielma toteutettiin kvantitatiivisin menetelmin korrelaatio- ja regressioanalyysien keinoin. Aineistona käytettiin paikkatieto-ohjelmalla käsiteltyjä Tilastokeskuksen tuottamia yhdyskuntarakenteen seuranta-aineistoja ja Oulun kaupungilta saatuja yläkoulujen päättökeskiarvoja. Tutkimuksessa pyrittiin löytämään vastaus tutkimuskysymykseen: Onko Oulun yläkouluissa näkyvissä päättökeskiarvojen eriytymistä, jota voitaisiin selittää koulun lähialueen sosioekonomisilla tekijöillä? Lisäksi tarkasteltiin mitkä analysoidut sosioekonomiset tekijät selittävät eriytymistä parhaiten, sekä kuinka tulokset vertautuvat aiempiin tuloksiin erityisesti Suomessa. Tutkimuksen teoreettinen viitekehys taustoittaa koulujen eriytymistä ja sen taustalla vaikuttavia ilmiöitä Suomessa.

Tulokset osoittivat, että Oulun yläkoulujen välillä on nähtävissä eriytymistä, joka on yhteydessä niiden lähialueiden sosioekonomisiin tekijöihin. Voimakkaimmin päättökeskiarvoihin yhteydessä oli keskiasteen koulutuksen saaneiden työssäkäyvien osuus koulun lähialueen työssäkäyvistä. Yhteys oli negatiivinen ja regressioanalyysien mukaan osuus selittää yksin noin kolmanneksen päättökeskiarvojen vaihtelusta. Tuloksiin liittyy joitakin epävarmuuksia, eikä niitä voida pitää kaikilta osin luotettavina tai merkitsevinä. Tulokset ovat samankaltaisia aiempien tutkimuksien kanssa, ja taustateoriaan nojautuen alueellisilla sosioekonomisilla tekijöillä voidaan selittää koulujen päättökeskiarvojen vaihtelua Oulussa.

Avainsanat: oppimistulokset, sosioekonomiset tekijät, alueellinen eriytyminen, segregatio, koulut, eriytyminen, tasa-arvo

Sisällys

1	Johdanto.....	4
1.1	Tutkimuksen tavoite.....	6
1.1.1	<i>Tutkimuskysymykset.....</i>	<i>8</i>
2	Koulujen eriytyminen Suomessa	9
2.1	Koulujen oppimistulosten eriytyminen.....	10
2.2	Perhetaustan yhteys oppimistuloksiin.....	13
2.3	Kouluvalinta koulujen eriytymisen mahdollistajana	16
2.4	Yhteiskunnallinen polarisaatio, alueellinen eriytyminen ja koulut.....	20
3	Aineisto ja metodologia	26
3.1	Aineiston kuvaus.....	28
3.2	Aineiston käsittely	32
3.2.1	<i>Lähialue-estimaatit ja tietojen poimiminen</i>	<i>34</i>
3.2.2	<i>Muuttujien valinta ja muodostaminen</i>	<i>42</i>
3.3	Analyysimenetelmät	45
3.3.1	<i>Korrelaatioanalyysi</i>	<i>46</i>
3.3.2	<i>Regressioanalyysi</i>	<i>49</i>
4	Analyysi ja tulokset.....	56
4.1	Yläkoulujen lähialue-estimaatit ilman oppilasalueita	56
4.1.1	<i>Työväestön koulutus ja oppimistulokset.....</i>	<i>57</i>
4.1.2	<i>Lähialueen tulot ja oppimistulokset</i>	<i>64</i>
4.1.3	<i>Regressioanalyysien tulokset</i>	<i>67</i>
4.2	Yläkoulujen lähialue-estimaatit oppilasalueiden kanssa	75
4.2.1	<i>Työväestön koulutus ja oppimistulokset.....</i>	<i>76</i>
4.2.2	<i>Lähialueen tulot ja oppimistulokset</i>	<i>78</i>
4.2.3	<i>Regressioanalyysien tulokset</i>	<i>81</i>
5	Tulosten koonti ja johtopäätökset	87
6	Eettiset kysymykset ja tutkimuksen luotettavuus.....	93
7	Pohdintaa	98
	Lähteet	103

1 Johdanto

Suomalaisen peruskoulujärjestelmän voidaan sanoa edustavan ja edistävän *mahdollisuuksien tasa-arvoa* sen perustavinta laatua olevassa muodossa. Täällä kaikki lapset ovat saaneet sosiaalisesta taustastaan, sukupuolestaan tai vaikkapa asuinalueestaan riippumatta mahdollisuuden osallistua koulutukseen ja menestyä siellä kykyjensä mukaisesti. (Silvennoinen, Kalalahti, & Varjo, 2016, 28–29.) Peruskoulu onkin pystynyt melko hyvin kompensoimaan oppilaiden erilaisia lähtökohtia ja tasaamaan eroja eri yhteiskuntaluokkien ja -ryhmien välillä. Suomalainen peruskoulujärjestelmä onkin ollut erinomainen sosioekonomisesta taustasta johtuvien erojen tasoittaja jo pitkään, mutta järjestelmä ei kuitenkaan ole täydellinen.

Mediassa on viime vuosina noussut aika ajoin esille, että koulujen tuottamat oppimistulokset ovat muuttuneet keskinäisessä vertailussa suuremmin poikkeaviksi. Mediassa kehitys on esitetty seuraavalla tavalla: suurissa kaupungeissa niin sanottujen parhaiden koulujen oppimistulokset paranevat, heikommassa puolestaan laskevat. Oppimistuloksien eriytymisen syynä ei kuitenkaan nähdä opettajien tai tutkijoiden mukaan eroja opettamisessa, vaan syyksi nähdään oppilasaineiksessa olevat erot. Yhteiskunnallisen eriarvoistumiskehityksen nähdään heijastuvan kouluihin oppilasaineuksen eriytymisen kautta. (Marttinen, 2019; Rossi, 2018; Vanas, 2020.)

Julkisessa keskustelussa esille ovat nousseet asuinalueiden sosioekonomisen eriytymiskehityksen vaikutukset kouluihin. Esille on nostettu esimerkiksi koulujen etnisen eriytymisen ja joidenkin koulujen maahanmuuttajalasten suurien määrien tuomat haasteet, sekä yleisesti huono-osaisuuden kasautuminen tietyille asuinalueille. Huono-osaisuus näyttäytyy myös näiden alueiden kouluissa erilaisina haasteina. (esim. Elo, 2018; Lepistö, 2019; Marttinen, 2019; Rossi, 2018.) Opettajien mukaan erot koulujen arkitodellisuudessa ovat hyvin suuria ja heijastuvat opetukseen. Joillakin alueilla sijaitsevilla kouluilla opettajien kertoman mukaan ei esimerkiksi kannata opettaa maanantaisin vaikeita asioita, sillä lapset eivät välttämättä ole saaneet viikonloppuna ruokaa (Rossi, 2018).

Huono- ja hyväosaisuuden kasautumisen tietyille asuinalueille ohella – ja siihen liittyen – mediassa on kannettu paljon huolta erityisesti koulutettujen vanhempien harrastamasta ”koulushoppailusta”, jossa vanhemmat pyrkivät välttämään lastensa päätymistä tiettyihin

kouluihin ja suosivat toisia kouluja. (Lepistö, 2019; Marttinen, 2019.) Vanhemmilla on uutisoinnin mukaan monia keinoja tähän: esimerkiksi asuinpaikan valinnalla, harvinaisen A-kielen opiskelulla ja painotetulle luokalle hakeutumisella voidaan vaikuttaa siihen missä lapsi käy kouluaan (Marttinen, 2019). Siinä missä vanhemmat välttelevät toisia kouluja, jotkin koulut ovat houkutelleet perheitä jopa niin paljon, että kaikkia halukkaita ja pääsykokeen läpäisseitä lapsia ei voida edes ottaa sisään kouluun tilojen riittämättömyyden vuoksi (Riihola, 2017). Erityisesti keskiluokkaisten, koulutettujen ja muuten hyvässä sosioekonomisessa asemassa olevien vanhempien ja huoltajien tekemien valintojen mahdollista suurta vaikutusta koulujen eriytymiskehityksen suunnannäyttäjänä on käsitelty mediassa paljon. (Johansson, 2019; Kuuskoski, 2019; Marttinen, 2019; Rautio, 2019; Rossi, 2018; Vanas, 2020; Vilka & Bernelius, 2019.)

Oman mausteensa koulujen eriytymisen soppaan tuo myös keväällä 2020 iskenyt COVID-19-viruksen aiheuttama pandemia. Jo aikaisin keväällä mediassa nousi esille mielipiteitä, joiden mukaan kyseisen koronaviruksen aiheuttama koulujen sulkeminen tulee näyttäytymään yhä suurempana epätasa-arvona koulumaailmassa. Esille nostettiin, että oppilaiden erilaiset tilanteet kotona tulevat paisuttamaan myös koulukohtaisia eroja (esim. Kauppinen, 2020). Jo kesän alussa paljastuikin, että eroja koulumenestyksessä todellakin nousi esiin ja niitä selitti monissa tapauksissa pitkälti perhetausta, kuten esimerkiksi Yleisradio (Eskonen, de Fresnes, & Pietarinen, 2020) uutisoi kesäkuun puolessa välissä. Samaisessa uutisessa puhuttiin myös siitä, kuinka koulussa pärjäämisen erot hyvä- ja heikko-osaisten välillä ovat kasvaneet jo jonkin aikaa Suomessa, ja kuinka perhetausta vaikuttaa merkittävästi siihen, miten pitkälle lapsi kouluttautuu. Monet tutkimustulokset vahvistavat tämän.

Hyvätuloisten lapset pärjäsivät etäkoulussa yleisesti ottaen paremmin erinäisistä syistä, jotka kytkeytyvät esimerkiksi perheen arvoihin ja kulttuuriin, mutta myös mahdollisuuksiin. Kevään ja kesän aikana käytiin keskustelua julkisesti esimerkiksi siitä, kuinka vanhempien työpaikalla ja työn laadulla on merkitystä sen suhteen, kuinka paljon he pystyvät osallistumaan ja auttamaan lastaan etäkoulussa. (esim. Eskonen ym., 2020; Kauppinen, 2020; Vaalavuo, 2020a.) Eläketurvakeskuksen tutkijan, dosentti Susan Kuivalaisen (2020) mukaan esimerkiksi kodin välineistössä, vanhempien kyvyissä ja mahdollisuuksissa auttaa koulutehtävissä, sekä aivan perustoimeentulossa on hyvin suuria eroja, jotka heijastuvat koulunkäyntiin erityisesti koronapandemian aikana, kun useat lapset eivät pääse päiviksi kouluun opiskelemaan. Koulu esimerkiksi saattaa olla jollekin lapselle ainoa paikka missä hän saa lämmintä ruokaa, joiltain lapsilta saattoi joissain kodeissa puuttua esimerkiksi turvallinen paikka opiskella tai oleellinen

tuki koulutehtävien suorittamiseen (Näveri, 2020; Vaalavuo, 2020a). Kuivalaisen (2020) mukaan lapsen mahdollisuudet pärjätä linkittyvät vahvasti vanhempien työmarkkina-asemaan, koulutustasoon, taloudellisiin resursseihin ja elämänhallintaan.

Koulujen eriytyminen on siis ajankohtainen aihe monellakin tapaa. Aiheesta käytiin keskustelua jo ennen pandemiaa, mutta pandemian jälkeen pelko epätasa-arvoisuuden kasvusta on nostanut entistäkin vahvemmin tapetille myös koulut ja niiden keskinäisen eriytymisen. Jotta suomalaisen peruskoulun ylpeyden aiheen oleva tasa-arvoisuus saataisiin turvattua, on syytä tutkia monia tasa-arvoon liittyviä aiheita, pyrkiä tunnistamaan ongelmia ja korjaamaan niitä. Koulujen eriytymisen voi katsoa olevan merkittävä haaste peruskoulujärjestelmän tasa-arvoisuudelle, sillä koulujen eriytyessä lasten tasavertaiset mahdollisuudet pärjätä koulutusinstituutiossamme ja myöhemmin koulutusta arvostavassa yhteiskunnassamme ovat uhattuina.

Tässä pro gradu -tutkielmassa kiinnostus kohdistuu erityisesti koulujen eriytymiseen alueellisten sosioekonomisten tekijöiden näkökulmasta. Kuten mediassa käydystä keskustelusta voidaan päätellä, on alueellinen eriarvoistuminen merkittävässä osassa koulujen eriytymiskehityksessä. Seuraavaksi esitellään vielä hieman tarkemmin tutkimuksen taustoja ja tavoitteita.

1.1 Tutkimuksen tavoite

Koulujen alueellista eriytymistä koskevia tutkimuksia on tehty Suomessa jonkin verran. Koulujen eriytymistä ja kouluvalintaa, sekä aiheita muuten sivuavia tutkimuksia on tehty erityisesti Helsinkiä käsitellen, vaikkakin joitakin tutkimuksia on tehty myös Espoossa, Vantaalla ja Turussa. Tutkimus on keskittynyt melko voimakkaasti Etelä-Suomeen ja erityisesti pääkaupunkiseudulle. Pohjois-Suomessa koulujen eriytymistä ja esimerkiksi kouluvalintoja käsittelevää tutkimusta ei ole tietääkseni tehty. Tämä pro gradu -tutkielma pyrkii osaltaan täyttämään tätä aukkoa tutkimuskentällä käsittelemällä koulujen oppimistulosten eriytymistä alueellisten sosioekonomisten tekijöiden näkökulmasta Pohjois-Suomessa.

Pohjois-Suomessa, tässä tapauksessa Oulussa, tehtävän tutkielman tärkeys perustuu muun muassa siihen, että pääkaupunkiseutu ja vielä suurempana kokonaisuutena Etelä-Suomi ovat hyvin erilaisia alueita verrattuna Suomen pohjoisempiin ja harvemmin asuttuihin alueisiin. On

mielestäni perusteltua väittää, että erityisesti pääkaupunkiseutu muodostaa kuplan, jossa saatuja tutkimustuloksia ei voida suoraan yleistää koskemaan muuta Suomea, eikä etenään Pohjois-Suomea.

Vaikka myös Oulu on yksi niistä harvoista kaupungeista, jotka ovat Suomessa muuttovoittoisia (kts. esim. Oulun kaupunki, 2020; Tilastokeskus, n.d.; Ylipiessa, 2018) on pääkaupunkiseutu kuitenkin kaupungistuvassakin Suomessa erityistapaus kokonsa ja muun vetovoimansa vuoksi. Pääkaupunkiseudun vertailu muuhun Suomeen on tutkimuksellisesti perusteltua monin tavoin, minkä lisäksi erityistä tutkimuksellista lisähyötyä voidaan saada vertaamalla sitä monin tavoin täysin poikkeavaan ympäristöön, Pohjois-Suomeen. Kärjistettynä voidaan esittää esimerkiksi, että pääkaupunkiseudulla etäisyydet ovat lähes aina lyhyempiä kuin Pohjois-Suomessa, minkä lisäksi elinkeinot, palvelut ja monet muut asiat poikkeavat vahvasti vaikkapa Oulun ja Helsingin välillä.

Helsinki on ylivoimaisesti Suomen suurin kaupunki, minkä lisäksi jotkin sen lähikaupungit – kuten Espoo ja Vantaa – ovat käytännössä kasvaneet siihen kiinni, ja ovat itsekkin Suomen suurimpien kaupunkien joukossa kärkisijoilla. Oulu puolestaan on alueellaan ainut suuri kaupunki, vaikkakin Kempele kasvaa hyvää vauhtia aivan Oulun kupeessa. Muut lähikunnat ovat kuitenkin asukasluvultaan pieniä ja harvaan asuttuja, mikä poikkeaa kovasti pääkaupunkiseudusta. Voidaan olettaa, että Etelä- ja Pohjois-Suomen kaupunkien väliset erot esimerkiksi juuri lähialueiden ja -kaupunkien, etäisyyksien, asukaslukujen ja muiden seikkojen erot heijastuvat myös kaupunkien kouluihin. Tästä syystä on mielekäästä tehdä tutkimusta myös Oulussa, sillä voidaan katsoa, että Helsingistä ja muista tiheästi asutetun Etelä-Suomen suurista kaupungeista saadut tutkimustulokset eivät sellaisinaan ole välttämättä suoraan yleistettävissä harvaan asuttuun Pohjois-Suomeen.

Oulu nähdään usein Pohjois-Suomen pääkaupunkina, sillä se on selvästi sen suurin kaupunki, ollen samalla merkittävä teollisuuskaupunki koko Suomen mittapuulla. Oulu sijaitsee karkeasti katsottuna noin puolessa välissä Manner-Suomea pohjois-etelä -suunnassa tarkasteltuna, mistä osoituksena Puolangassa, noin sata kilometriä Oulusta lähes suoraan itään sijaitseva Manner-Suomen keskipiste. Näin ollen voidaan katsoa, että raja Pohjois- ja Etelä-Suomen välillä kulkee jokseenkin Oulun kohdalla, mikäli maa jaetaan kahteen osaan. Pohjois-Suomelle ei ole olemassa mitään vakiintunutta määritelmää, mutta tässä pro gradu -tutkielmassa sillä tarkoitetaan Pohjois-Suomen aluehallintaviraston aluetta (kts. Aluehallintavirasto, 2016), sekä siitä pohjoiseen sijaitsevia alueita. Näin ollen myös Lapin aluehallintavirastolle kuuluvat alueet

käsitetään tässä yhteydessä kuuluvaksi Pohjois-Suomeen. Pohjois-Suomen aluehallintaviraston alue voidaan näin nähdä myös Pohjois-Suomen eteläosana, joka pitää sisällään Pohjois-Pohjanmaan ja Kainuun maakunnat, pohjoisosan muodostuessa Lapin maakunnasta.

On tärkeää huomata, että Pohjois-Suomi on määritelmästä riippumatta joka tapauksessa hyvin laaja maa-alue, johon mahtuu sekä muutamia Suomen mittapuulla suuria tai suurehkoja kaupunkeja, että laajoja hyvin harvaan asuttuja alueita. Tutkimuksessa keskitytään erityisesti Ouluun, eikä tutkimuksen tuloksia voida näin ollen yleistää koskemaan koko Pohjois-Suomea. Tämä ei myöskään ole tutkimuksen tarkoituksena, vaan tutkimuksen tarkoituksena on selvittää, näkyykö nimenomaan Oulun kouluissa mahdollisesti alueellisella sosioekonomisella eriytymisellä selitettävää yläkoulujen oppimistulosten eriytymistä. Oulun voidaan kuitenkin epäilemättä katsoa kuuluvan Pohjois-Suomeen, ja näin ollen Oulussa tehtävä tutkimus vastaa osaltaan Pohjois-Suomessa tehdyn aiheeseen liittyvän tutkimuksen puutteeseen, tasapainottaen aavistuksen Etelä- ja Pohjois-Suomen välistä epäsuhdannetta tämän tutkimusaiheen osalta.

Tutkimuksessa kiinnostuksen kohteena ovat Oulun kaupungin yläkoulut ja niitä ympäröivät lähialueet tai naapurustot. Tavoitteena on selvittää, onko yläkoulujen lähialueiden sosioekonomisilla tekijöillä yhteyksiä yläkoulujen oppimistuloksiin. Tutkimuksessa tarkastelu keskittyy alueella asuvan työssäkäyvän väestön koulutustasoon sekä alueen yleiseen tulotasoon ja niiden yhteyteen koulujen oppimistuloksiin. Edellä kuvattujen sosioekonomisten tekijöiden tärkeyttä tutkimusaiheen kannalta avataan tarkemmin myöhemmin.

1.1.1 Tutkimuskysymykset

Tutkimuskysymyksiksi ovat taustateorian ja aiempien tutkimusten tuntemuksen perusteella muotoutuneet:

- Onko Oulun yläkouluissa näkyvissä päättökeskiarvojen eriytymistä, jota voitaisiin selittää koulun lähialueen sosioekonomisilla tekijöillä?
 - Jos eriytymistä on havaittavissa, mitkä tutkimuksessa analysoidut tekijät selittävät sitä parhaiten?
 - Kuinka tulokset vertautuvat aiempiin tuloksiin Suomesta?

2 Koulujen eriytyminen Suomessa

Koulujen eriytyminen on ilmiö, jota voidaan tarkastella monelta eri kannalta. Jo pelkästään siitä, mitä pidetään eriytymisenä, voidaan olla monta mieltä. Koulujen eriytyminen voi tarkoittaa esimerkiksi oppimistulosten eriytymistä, oppiaineiden erilaisia painotuksia koulujen välillä, oppilaspohjan eriytymistä tai resurssien erilaisia määriä erilaisilla kouluilla. Eriytyminen voidaan siis käsittää monella tavalla, eikä ole myöskään yksiselitteistä, että kaikki eriytyminen olisi lähtökohtaisesti huonoa tai negatiivista. Esimerkiksi uusliberalistinen koulutuspolitiikka nojaa vahvasti siihen, että yksilöille tarjotaan mahdollisuuksia hyödyntää omia kykyjään ja taipumuksiaan mahdollisimman tehokkaasti. (Silvennoinen ym., 2016, 30–31)

Koulujen eriytymisestä puhuttaessa onkin syytä tarkentaa, millaista eriytymistä käsitellään: Ollaanko esimerkiksi kiinnostuneita siitä, millaisia oppilaita kouluissa opiskelee vai siitä millaisia oppimistuloksia koulu tuottaa? Toisaalta kiinnostuneita voidaan olla myös esimerkiksi siitä, millaiset resurssit eri kouluilla on käytössään ja miten ne eroavat, sekä miten erilaiset resurssit vaikuttavat koulun toimintaan ja tuloksiin.

Jos eriytymistä on vaikea määritellä suoraan, on kenties vielä vaikeampaa hahmottaa sitä, mistä eriytyminen johtuu ja mitä siitä seuraa. Koulun toimintaan, prosesseihin ja suorituksiin vaikuttavat nimittäin lukuisat kontekstit ja yhteydet. Näitä ovat esimerkiksi muut kuin sosioekonomisessa taustassa näkyvät erot oppilasarakenteessa, kuten oppilaiden etninen tausta, turvapaikanhakijataustaisten lasten osuus koulun oppilaista ja erityistarpeita omaavien lasten osuus, sekä alue- ja koulukohtaiset muut piirteet sisältäen esimerkiksi sen, onko kyseessä maalais- tai kaupunkikoulu, mikä on koulun asema suhteessa alueen muihin kouluihin, sekä paikallinen koulutuspolitiikka. (Thrupp, 2018, 50–51.)

Tässä luvussa pyritään käsittelemään koulujen eriytymistä Suomessa esittelemällä sitä, kuinka koulujen eriytyminen havaitaan Suomessa. Lisäksi pyritään esittelemään joitakin tekijöitä, jotka tutkimusten perusteella vaikuttavat koulujen eriytymiskehityksen etenemiseen Suomessa. Kyseessä on hyvin laaja yhteiskunnallinen ja sosiaalinen ilmiö, jonka kohdalla on mahdotonta pro gradu -tutkielman laajuudessa käsitellä kaikkia mahdollisia selityksiä ja näkökulmia. Tässä yhteydessä pyritäänkin tuomaan esille niitä puolia ilmiöstä, jotka kaikista selvimmin liittyvät tutkimuksen aiheeseen ja sen tulkintaan.

2.1 Koulujen oppimistulosten eriytyminen

Puhuttaessa koulujen eriytymisestä varmasti ensimmäisenä suurimmalla osalla mieleen nousee ajatus siitä, että koulut ovat eri tasoisia. Tämä näyttäytyy ennen kaikkea sitä kautta, millaisia oppimistuloksia kouluissa pystytään tuottamaan. Jo pitkään on kiinnitetty huomiota siihen, että oppimistulokset eri koulujen väleillä ovat kasvaneet. Suomessa on nähtävissä kehityskulkua, jossa eri koulujen oppimistulokset poikkeavat vuosi vuodelta yhä voimakkaammin toisistaan jopa yksittäisen kaupungin sisällä. Eriytyminen näytti ainakin välillä olevan erityisen vahvaa Helsingissä, jossa koulujen väliset erot oppimistuloksissa (18 %) olivat samaa luokkaa kuin esimerkiksi Azerbaidzanissa, Irlannissa ja Kanadassa. (Bernelius, 2011, 479; Bernelius, 2013a, 17.) Koko Suomen tasolla koulujen väliset erot ovat olleet tyypillisesti huomattavasti tätä pienempiä. Koko maan tasolla koulujen tuottama selitysosuus koulujen oppimistulosten eroista on vaihdellut eri tutkimusten mukaan välillä 5–9 %, eli yksin Helsingin koulujen tuottama selitysosuus 18 % on jopa kaksin- tai kolminkertainen muuhun maahan verrattuna. (Bernelius, 2013a, 67 ja 80.) Myös tuoreimmat PISA 2018 -tulokset kertovat siitä, että pääkaupunkiseudulla koulujen välinen vaihtelu olisi muuta Suomea suurempaa ja että parhaat ja heikoimmat koulut Suomessa ovat pääkaupunkiseudulla (Leino ym., 2019, 62).

Hieman tuoreemmassa tutkimuksessa Helsingin yläkoulujen selitysosuudeksi yhdeksäsluokkalaisten oppimistuloksista saatiin 8,7 % kun oppilaiden lähtötasoa ei huomioitu. Kun kyseisten oppilaiden seitsemännen luokan lähtötaso otettiin huomioon sekä koulu-, että yksilötasolla, koulun selitysosuudeksi saatiin enää hieman yli 3 % oppimistulosten vaihtelusta. (Vainikainen ym., 2016, 220 ja 224.) Mielenkiintoisena sivukaneettina mainittakoon, että tutkittaessa PISA 2009 -vuoden tuloksia Suomen ruotsinkielisten koulujen osalta, havaittiin, että pienet koulut tuottavat tilastollisesti merkittävästi heikompia oppimistuloksia kuin suuret. Eräs selittävä tekijä voi olla pienempien koulujen keskimääräistä matalammat oppilaiden sosioekonomiset taustat, jotka johtuvat mahdollisesti siitä, että pienet koulut sijaitsevat usein maalla. Menemättä tarkemmin yksityiskohtiin, maalla asuvien perheiden sosioekonomiset taustatekijät ovat usein pienemmät kuin kaupungissa asuvien. (Harju-Luukkainen, Sulkunen, & Vettenranta, 2015, 67–69.)

Kansainvälisesti vertailtuna koulujen väliset erot ovat Suomessa olleet aina pieniä riippumatta siitä onko tarkastelussa käytetty matematiikkaa, lukutaitoa vai luonnontieteiden osaamista. Samalla tasolla Suomen kanssa on ollut ainakin Norja, muiden Pohjoismaiden osalta vaihtelu on ollut hieman suurempaa. Erot ovat Suomessa kuitenkin olleet lievässä kasvussa. Myös

tuoreimmissa PISA-tuloksissa näyttää siltä, että suhteutettuna muihin OECD-maihin Suomen koulujen väliset erot ovat edelleen pieniä. Suomessa koulujen välinen vaihtelu lukutaidossa oli vuoden 2018 PISAssa 7 % kun keskimääräinen vaihtelu OECD-maissa oli 29 prosenttiyksikköä. Peruskoulujen välinen vaihtelu oli vuonna 2018 myös pienempää kuin esimerkiksi vuosina 2012 ja 2015. Tulokset pysyvät samoina tarkastellessa eroja matematiikan tai luonnontieteiden kautta, eikä näin ollen PISA-tulosten perusteella eriytymiskehitystä voida pitää erityisen hälyttävänä. (Leino ym., 57–59.) Toisaalta, jos tarkastelua muutetaan hieman, ja tarkkaillaankin suomalaiskoulujen tuloksia siten, että ne on jaettu kymmenyksiin oppilaidensa keskimääräisten PISA-tulosten mukaan, eroja näyttäisi syntyvän parhaiden ja heikoimpien koulujen välille, ja erot ovat todellakin kasvussa. Erot havaittiin jo vuoden 2015 PISA-tuloksissa, jolloin Suomi myös menetti parhaiten ja heikoiten menestyvien koulujen välisen pienimmän eron tittelinsä Islannille. Erojen kasvu vaikuttaa jatkuneen tuoreimmissa tuloksissa. Tässä tarkastelussa erot johtuvat siitä, että heikoimpien koulujen tulokset ovat vuosi vuodelta heikompia. Hyvien koulujen tuloksissa sen sijaan ei näy merkittäviä muutoksia. Suurin osa kouluista kuitenkin sijoittuu jakauman keskivaiheille, jossa vaihtelu on lähes merkityksettömän pientä. (Vettenranta ym. 2016, 57–59; Leino ym., 2019, 59–60.)

Asuinalueiden välisten erojen kasvu ja sen yhteys koulujen oppimistulosten eriytymiseen on todettu Suomessa useissa tutkimuksissa. Oppilaiden oppimistulokset ovat ainakin Helsingin kouluissa selvästi yhteydessä koulujen oppilaspohjien yleisiin piirteisiin, korostuen yläkoulujen kohdalla suhteessa alakouluihin. Helsingin kaupunginosien ja koulujen eriytyminen on hyvin samankaltaista kuin useissa muissa eurooppalaisissa ja pohjoisamerikkalaisissa kaupungeissa. (Bernelius, 2013a, 67 ja 80). Koulujen osaamiserojen kasvu kytkeytyy Helsingissä vahvasti kaupunginosien segregaatioon. Kehitystä kiihdyttävät muuttoliike pois tietyiltä alueilta ja kouluvalinnat. (From ym., 2014, 553.) Verrattaessa helsinkiläisten 7.-luokkaisten oppimistuloksia, olivat muuhun kuin lähikouluunsa hakeutuneiden oppilaiden oppimistulokset parempia kuin vertailuryhmällä. Erot lähikouluunsa jäävien ja muualle hakeutuvien seitsemäsluokkalaisten tuloksissa ovat jopa kolmannes tulosten välisestä keskihajonnasta, mutta eivät kuitenkaan selity pelkällä koulun vaikutuksella. (Bernelius, 2013a, 67.) Suomalaisessa muodossa kouluvalinnasta – joka siis toimii vahvasti painotetun opetuksen ryhmiin valikoitumisen kautta – vaikuttaa korostuvan koulutuksen periytyvyys ja vanhempien kulttuurisen pääoman merkitys. (Bernelius, 2013a, 68; Kalalahti, Silvennoinen, & Varjo, 2015a, 23.)

Koulujen lähialueiden eriytymisen tuottamia vaihteluita oppimistulosten suhteen on kuitenkin tarkasteltu Suomessa melko vähän. Pioneerina aiheeseen on toiminut Bernelius. Jo Berneliuksen pro gradu -tutkielman (2005) tulokset osoittivat, että asuinalueiden väestörakenteen eriytymisellä on yhteys koulujen oppimistuloksiin. Vaikka alueellisten sosioekonomisten tekijöiden yhteyksiä oppimistuloksiin on tutkinut Suomessa lähinnä Bernelius, on myös PISA-tuloksista tehty tarkasteluja koulun sijainnin perusteella. Tarkastelut ovat kuitenkin rajoittuneet lähinnä kaupunkikokoja tarkastelevalle tasolle. PISA 2018 -tuloksissa havaittiin, että mitä pienemmällä paikkakunnalla koulu sijaitsi, sitä heikompaa oli poikien osaaminen. Tyttöjen osaamiseen paikkakunnan koolla, eli koulun sijaintiympäristöllä ei kuitenkaan ollut vaikutusta. (Leino ym., 2019, 67–68.)

Väestön sosioekonomista ja etnistä rakennetta kuvaavien muuttujien avulla pystytään Berneliuksen mukaan rakentamaan malli, joka selittää tilastollisesti merkitsevästi jopa 50 % yläkoulujen oppimistulosten vaihtelusta. (Bernelius, 2013a, 18.) Kuusela kuitenkin esittää, että yhtä muuttujaa – äidin koulutustasoa – käyttämällä voidaan päästä paljon korkeampiin tuloksiin; Espoossa selitysaste lähentelee 90 %, Helsingissä 80 %, mutta Vantaalla vain 30 %. Kuuselan mukaan tämä johtuu siitä, että Vantaalla eri koulujen oppilaiden äitien koulutustasot vaihtelivat huomattavasti vähemmän kuin Espoossa ja Helsingissä. (Kuusela, 2010, 46–47.) Myös Ruotsissa on saatu samoihin aikoihin tuloksia, joiden mukaan koulujen etninen eriytyminen ja perheiden, sekä oppilaiden ja koulujen naapurustojen sosioekonomiset taustat selittävät vahvasti saavutettuja oppimistuloksia (Szulkin & Jonsson, 2006). Hieman tuoreempi vertailu maiden pääkaupunkien välillä osoittaa, että koulujen oppimistuloksien väliset erot Tukholmassa olivat hieman suurempia kuin Helsingissä. Lisäksi on saatu selville, että ainakin jotkin perhetaustaan liittyvät tekijät, kuten vanhempien työttömyys ja perherakenne, ovat Tukholmassa Helsinkiä vahvemmin yhteydessä oppilaiden oppimistuloksiin. (Modin, Karvonen, Rahkonen, & Östberg, 2015, 473–480.)

Helsingissä on kokeiltu sosioekonomisesti heikoimpien alueiden koulujen kohdalla niille kohdennettua erityistä tukea, jonka tulokset ovat olleet positiivisia. Tämä niin kutsuttu positiivisen diskriminaation politiikka on tukenut oppimista, vähentänyt koulupudokkaiden määrää koulutuspolun taitekohdissa sekä parantanut oppilaiden mahdollisuuksia jatko-opintoihin erityisesti maahanmuuttajataustaisiin oppilaisiin kohdennettuna. (Nevalainen, 2018, 4; Silliman, 2017, 39–40.) Tulokset ovat lisäksi olleet pysyviä, mutta sikäli ainutlaatuisia, että samanlaisia tuloksia ei ole saatu muissa Suomen suurissa kaupungeissa (Silliman, 2017, 44–45). Kouluihin kohdistettu tuki ei kuitenkaan yksin riitä, vaan se tarvitsee rinnalleen myös

muita toimenpiteitä, jotka pitkäjänteisesti ja määrätietoisesti parantavat alueen aikuisten mahdollisuuksia. Toimenpiteitä voivat olla esimerkiksi työikäisten koulutuksen tukeminen ja maahanmuuttajien kielitaidon kohentaminen. Tällaiset toimenpiteet vaikuttavat pitkällä aikajänteellä koulujen eriytymisen juurisyihin, niihin asioihin, jotka luovat alueellista segregatiota. (Nevalainen, 2018, 4.) Fromin ja kumppanien (2014, 553–554) mukaan koulujen eriytymistä voidaan vastustaa takaamalla kaikille asuinalueille laadukkaat lähiperuskoulut, joissa eritaustaiset oppilaat sekoittuvat luontevasti, sekä ehkäisemällä ja pysäyttämällä oppimistulosten eriytymistä muillakin keinoin. Ruotsissa on todettu, että koulujen yleisen tason nostamiseksi paras keino lienee lisätä maahanmuuttajien työllisyysastetta ja taloudellista tilannetta (Szulkin & Jonsson, 2006, 25).

Koulujen oppimistulokset ovat siis Suomessa eriytyneet viimeisten vuosikymmenten aikana jatkuvasti. Erot eivät kuitenkaan ole kovin dramaattisia koko maan tasolla, mutta jos tarkastellaan pelkästään pääkaupunkiseutua, koulujen väliset erot näyttäytyvät paljon vahvempina. Muihin maihin verrattuna Suomessa tilanne koulujen oppimistulosten eriytymisen suhteen on vielä kohtuullisen hyvä, mutta erityisesti heikoimpien ja parhaimpien koulujen oppimistulosten eron kasvu aiheuttaa kuitenkin huolta. Suomessa on kokeiltu myös toimenpiteitä koulujen oppimistulosten eriytymisen ehkäisemiseksi, ja pääkaupunkiseudulla tulokset ovatkin olleet hyviä, vaikka niitä ei olekaan saatu muissa suurissa kaupungeissa toistettua. Koulujen oppimistulosten eriytymisen tutkimusta on syytä jatkaa, jotta nykyinen kehityskulku saadaan katkaistua tai sen suunta käännettyä.

2.2 Perhetaustan yhteys oppimistuloksiin

Tutkimukset ovat jo pitkään osoittaneet, että vanhempien sosioekonomisella taustalla on vaikutusta lasten oppimistuloksiin kaikissa yhteiskunnissa, mutta korrelaation voimakkuus vaihtelee. Suomessa yhteys vanhemman sosioekonomisen aseman ja lapsen oppimistulosten välillä on kansainvälisessä vertailussa pieni. Vuoden 2015 PISA-tulosten mukaan perhetaustan yhteys oppimistuloksiin oli kuitenkin enää keskiarvoa kansainvälisessä vertailussa, kun aiemmin se oli ollut erittäin vähäinen, mitä tietysti voidaan pitää hyvänä suorituksena. Sosioekonomisen taustan vaikutus onkin ollut nousussa jo pitkään, ja näkynyt erityisesti vanhempien koulutuksen kautta. (Lehti & Laaninen, 2020, 4; Salmela-Aro & Chmielewski, 2019, 7–12). Myös tuoreimmat PISA-tulokset osoittavat kehityskulun jatkuneen vastaavalla

tavalla (Leino ym., 49–53). Kansainvälisessä vertailussa kehitys Suomessa ei kuitenkaan ole ollut erityisen nopeaa (kts. esim. Chmielewski, 2019, 533).

Perhetaustan yhteys oppimistuloksiin vaikuttaisi siis näyttäytyvän erityisesti vanhempien koulutustaustan kautta. Vanhemman koulutuksella Suomessa on hieman hänen ammattiasemaansa vahvempi yhteys lapsen oppimistulosten selittäjänä. (Lehti & Laaninen, 2020, 14.) Erityisesti äidin koulutustasolla vaikuttaisi olevan vahva positiivinen yhteys lapsen kouluarvosanaan (esim. Kalalahti, 2014). Myös kaikissa viimeisimmissä PISA-tutkimuksissa on havaittu oppilaan sosioekonomisen taustan yhteys oppimistuloksiin (Välijärvi ym., 2015, 136; Vettenranta, 2016; Leino ym., 2019).

Tulokset siis osoittavat melko kiistatta oppimistulosten olevan yhteydessä perhetaustaan. Monimutkaisempaa onkin selittää, miten yhteys syntyy. Tiedetään kuitenkin, että vanhemmat siirtävät lapsilleen tietynlaisia arvoja, asenteita, tietoja ja taitoja. Ilmiötä on tavattu kutsua yhteiskuntatieteissä sosialisatioksi ja sosialisatioprosessissa, joka alkaa jo lapsen syntymästä, lapsen kanssa jaetaan yhteisön tärkeimmät arvot, normit, roolit, tavat ja muu sellainen, joka auttaa häntä toimimaan yhteisön jäsenenä. (Antikainen, Rinne, & Koski, 2013, 42–44). Sosialisatioprosessin voi katsoa olevan yksi tekijä perhetaustan ja oppimistulosten välisen korrelaation selittäjänä. Koulutuksen arvostus eri yhteiskuntaluokkiin kuuluvien perheiden välillä voi esimerkiksi olla hyvin erilainen, minkä vuoksi esimerkiksi korkeakoulutettujen vanhempien perheissä saatetaan koulumenestystä pitää yleisesti korkeammassa arvossa kuin matalammin koulutettujen. On esimerkiksi tutkittu, että akateemisten taitojen ja yleissivistävän koulutuksen arvostus ei ole kovin korkeaa työväestön keskuudessa (Silvennoinen, Rinne, Kalalahti, & Varjo, 2015, 443). Esimerkiksi Paul Willis (1984) päätteli uraauurtavassa tutkimuksessaan sosiaaliluokkien uudistumisesta, että sosiaalistuttavalla vertaisryhmällä on suuri merkitys sen suhteen, pidetäänkö koulutusta arvostettavana ja tavoittelemisen arvoisena.

Sosialisatioprosessi voi vaikuttaa koulumenestykseen monin tavoin, mutta esimerkiksi asenteiden siirtyminen vaikuttaa siten, että koulumyönteisten ja koulunkäynnistä pitävien lasten on tutkittu pärjäävän koulussa hieman paremmin kuin ei-myönteisten lasten. Koulumyönteiset lapset myös vaikuttaisivat saavan enemmän tukea koulunkäyntiinsä kuin vähemmän myönteiset lapset. Tukea saavien lasten vanhemmat puolestaan ovat tutkimusten mukaan yleisellä tasolla korkeammin koulutettuja kuin tukea saamattomien lasten. Vanhempien aktiivisuudella, kiinnostuneisuudella ja myönteisyydellä lapsen koulunkäyntiä kohtaan voidaan olettaa olevan positiivinen vaikutus lapsen koulumyönteisyyteen, mikä

puolestaan nostaa koulumenestyksen tasoa. Vastaavasti vanhempien koulukielteisyyden on havaittu heikentävän lapsen koulumenestystä muun muassa tehtäväorientoitumisen vähenemisen kautta. Positiiviset odotukset, kannustus ja hyvä itseluottamus saattavat johtaa jopa ylisuoriutumiseen koulussa, näiden puuttuminen ja vastakohtat puolestaan ennustavat alisuoriutumista olemassa oleviin kykyihin nähden. (Kalalahti, 2012, 381–386; Kalalahti, 2014, 49–51; Kärkkäinen, 2004, 90; Koivisto, 2018, 24–25.) Myönteinen suhtautuminen opiskeluun ja hyvä koulumenestys vaikuttavat olevan yhteydessä vanhempien hyvään tulotasoon, korkeaan koulutukseen ja matalaan työttömyyteen. Heikko koulumenestys ja kielteinen suhtautuminen koulunkäyntiä kohtaan puolestaan vanhempien matalaan ansiotasoon, korkeaan työttömyyteen ja vähäisempään koulutukseen. (Kuusela, 2010, 47.)

Perhetausta voi vaikuttaa oppimistuloksiin lukuisilla muillakin tavoilla. Esimerkiksi vanhempien koulutuskulttuurisella pääomalla voi olla suuri merkitys lapsen koulumenestyksen selittäjänä. Esimerkiksi oman koulutuksen puute voi haitata muuten aktiivisten, kannustavien ja koulumyönteisten vanhempien kykyä auttaa esimerkiksi läksyissä ja kokeisiin valmistautumisessa. (Silvennoinen ym., 2015, 443). Sen sijaan vanhemmat, jotka ovat itse kouluttautuneet korkealle, pystyvät hyvin suurella todennäköisyydellä auttamaan lastaan kotitehtävissä ja kokeisiin valmistautumisessa, mikä puolestaan näkyy koulumenestyksessä.

Perhetaustan vaikutus oppimistuloksiin selittyy siis pitkälti erilaisilla sosiaalisilla ja kulttuurisilla tekijöillä. Vanhempien koulumyönteisyys, kyky tukea, auttaa ja ohjata, sekä esimerkiksi sosialisointi kautta vanhemmilta lapsille siirtyvä koulutuksen ja koulun arvostaminen ovat tekijöitä, jotka osaltaan selittävät sitä, miksi erilaisista sosioekonomisista taustoista tulevat lapset menestyvät koulussa eri tavoin. Lisäksi on mahdollista, että esimerkiksi perheen varallisuudella on vaikkapa asuinalueen valinnan, ja siitä seuraavan koulun kautta yhteyttä lapsen koulumenestykseen. Voi esimerkiksi olla, että varakas perhe kykenee muuttamaan alueelle, jossa on tunnetusti hyvä koulu ja näin perhetausta vaikuttaa oppimistuloksiin myös taloudellisten resurssien kautta.

Erilaisten sosioekonomisten taustatekijöiden yhteyksiä ja vaikutusmekanismeja oppimistuloksiin voitaisiin käsitellä hyvinkin pitkästi. Tässä tutkimuksessa aiheeseen ei kuitenkaan uppouduta sen syvällisemmin, vaan tässä luvussa tahdottiin vain tuoda lukijan tietoon joitakin keskeisiä tuloksia perhetaustan ja oppimistulosten välillä, sekä esitellä joitakin mahdollisia vaikutusmekanismeja. Tärkeintä on tiedostaa, että perhetausta eli perheen sosioekonomiset tekijät ovat yhteydessä lasten oppimistuloksiin, ja että yhteydet voivat olla

hyvin monisyisiä. Lisäksi huomioitavaa on, että sosioekonomisista taustatekijöistä erityisesti vanhempien koulutustasolla vaikuttaa olevan positiivinen yhteys lapsen koulumenestykseen ja oppimistuloksiin.

2.3 Kouluvalinta koulujen eriytymisen mahdollistajana

Voidaan katsoa, että pohja nykyisen kaltaiselle koulujen eriytymiskehitykselle luotiin Suomessa 1980-luvulla, kun koulujen oppilasvalinnan sääntelyä alettiin purkamaan. Aiemmin kuntien koulupiirien olemassaolo oli ohjannut ensin lasten sijoittumista kansakouluihin, sen jälkeen 1970-luvulta alkaen peruskouluihin. Sitten säännöstelyä alettiin purkamaan 1980-luvulla, mikä johti 1990-luvulla koulujen perheiden itsensä tekemien kouluvalintojen tulemiseen osaksi suomalaista peruskoulujärjestelmää. Yksityiskohtaisesta opetussuunnitelmasta luovuttiin samoihin aikoihin, ja 1994 otettiin käyttöön ensimmäiset opetussuunnitelman perusteet. Opetussuunnitelman perusteet mahdollistivat osaltaan sen, että kouluissa alkoi esiintyä painotuseroja, sillä kunnat ja koulut alkoivat tehdä itse omat opetussuunnitelmansa. (Kalalahti, Silvennoinen, & Varjo, 2015b, 372–374; Seppänen, 2006, 9, 11–12 ja 62–71.) Opetussuunnitelman perusteet siis ehkä hieman yllättäenkin mahdollistivat myös eriytymiskehityksen, sen sijaan että ne olisivat vain lisänneet koulujen itsenäisyyttä opetuksen järjestämisen osalta. Vuodesta 1994 alkaen etenkin suurissa kaupungeissa on ollut mahdollista hakea oppilaaksi muuhun kuin kunnan osoittamaan lähikouluun. Samoihin aikoihin tapahtui myös toinen merkittävä muutos, kun erikoisluokkien määrä alkoi lisääntyä voimakkaasti. (Seppänen, 2006, 175.)

Kouluvalinnan vapautumisen eräänä mahdollistajana on nähty 1990-luvun laman aikoihin Suomeen omaksuttu uusliberalistinen hallintatapa. Uusliberalistisessa ideologiassa on perinteisesti tavattu ajatella, että parhaita tuloksia saadaan aikaan kilpailun avulla. Koulutuspolitiikan alueella tämä tarkoittaa esimerkiksi sitä, että koulut ja oppilaitokset teoriassa kilpailisivat parhaista opettajista ja oppilaista, jotka puolestaan kilpailisivat keskenään pääsystä parhaisiin kouluihin ja oppilaitoksiin. Kilpailussa jalkoihin jäävien koulujen, opettajien ja oppilaiden tarvitsee tällöin pyrkiä kehittämään omaa toimintaansa ja kenties keksiä jotain uutta, joka parantaisi näiden pärjäämistä kilpailussa parempia kouluja vastaan. Näin kilpailu tulisi ajan myötä parantaneeksi koulutuksen laatua kaikilla tasoilla optimoiden samalla myös oppilaiden kykyjen käyttöä ja yritteliäisyyttä. Jotta edellä mainittu skenaario voisi tapahtua, täytyy koulutusmarkkinoiden kuitenkin olla sellaisia, että ne tarjoavat

erilaisille yksilöille erilaisia mahdollisuuksia käyttää kykyjään, taipumuksiaan ja esimerkiksi erilaisia sosioekonomisia pääomiaan – joita ovat sosiaalinen, kulttuurillinen ja taloudellinen – parhaalla mahdollisella tavalla. Uusliberalistisen ajattelu- ja hallintatavan nousun vuoksi Suomessakin alettiin 1990-luvulla entistä vahvemmin ajamaan kouluvalinnan mahdollisuutta, mikä johti lopulta kouluvalinnan vapauttamiseen ja erilaisten koulutusmahdollisuuksien valinnan voimakkaaseen kasvuun vuosikymmenen puolivälistä eteenpäin. (Silvennoinen ym., 2016, 15–17 ja 30.)

Uusliberalistisen koulutuspolitiikan nousun vaikutuksia voidaanakin varmasti tulkita näkökulman mukaan sekä positiivisina että negatiivisina. Englannissa, mistä markkinalähtöinen uusliberalistinen koulutuspolitiikka on lähtöisin, voidaan sen todeta jopa epäonnistuneen, minkä vuoksi maassa suunniteltiin ainakin jollain tasolla jo vuosikymmen sitten paluuta perinteisempiin sosiaalidemokraattisiin lähestymistapoihin koulutuksen järjestämisen suhteen (Whitty, 2011, 34–35). Suomessa uusliberalistisen koulutuspolitiikka ei ole mennyt missään vaiheessa aivan yhtä pitkälle kuin Englannissa, mutta sen nousun voidaan katsoa osaltaan olevan vastuussa siitä, että peruskoulujärjestelmämme sisällä on alkanut esiintyä eriytymistä eri tasoilla. Koulutusjärjestelmämme voidaan katsoa muuttuneen uusliberalistisen koulutuspolitiikan myötä selektiivisemmäksi, monimuotoistuneemmaksi ja yksilöllistyneemmäksi. Samalla sen voidaan kuitenkin katsoa alkaneen eriytyä sisäisesti. Yhtenäiseen järjestelmään on syntynyt mekanismeja, jotka rakentavat esimerkiksi yhteiskuntaluokan, alueen, vähemmistöryhmän tai sukupuolen kautta eriytymistä. (From ym., 2014; Silvennoinen ym., 2016, 28–29; Whitty, 2011, 38–39.)

Suomessa suurin osa kouluista on edelleen yhteiskunnan rahoittamia julkisia kouluja, joten koulujen – ja toisaalta myös oppilaiden – välinen kilpailu rajoittuu pitkälti kouluvalinnan vapautumisen tuottamille *julkiskoulumarkkinoille* (Leivo, 2008, 36). Maassamme koulujen eriytymisen ilmiö onkin nykyisin kaikista selkeimmin nähtävissä siinä, millaisia erilaisia erityisluokkia peruskoulut tarjoavat. Esimerkiksi Koskela (2016, 120–121) esittää, että kouluvalintamahdollisuuksista hyvin perillä olevat vanhemmat eivät vie lastaan lähikouluun, vaan pyrkivät saamaan lapsensa ”paremmaksi koettuun” kouluun. He voivat pyrkiä saamaan lapsensa esimerkiksi kieli-, liikunta-, tai musiikkiluokalle, joita kaikki lähikoulut eivät pysty tarjoamaan. Painotetun opetuksen tarjonnasta onkin muotoutunut kunnallisia kouluvalintatiloja, kouluja ja lopulta oppilaita eriyttävä tekijä. Kunnilla on erilaisia mahdollisuuksia tarjota erilaisia painotuksia, mikä heijastuu niiden kouluihin. Kaikki Suomen suurimmat kunnat voivat esimerkiksi tarjota ensimmäiseltä luokalta alkavia kielipainotuksia,

sekä kolmannelta ja seitsemänneltä luokalta alkavia mahdollisuuksia painotettuun opetukseen tietyissä oppiaineissa. (Kalalahti ym., 2015a, 23.) Pienillä kunnilla todennäköisesti ei ole yhtä hyviä mahdollisuuksia tällaiseen, ja näin ollen jo pelkästään eri kunnissa sijaitsevia kouluja vertaamalla voidaan todeta niiden eriytyneen valintamahdollisuuksien kautta.

Painotettuun opetukseen vaikuttavat valikoituvan erityisesti korkeakoulutettujen perheiden lapset. Painotetun opetuksen ryhmissä on havaittu saavutettavan myös keskimääräistä parempia oppimistuloksia. Samanlaisia tuloksia on saatu myös muun kuin lähikoulunsa valitsevien oppilaiden kohdalla. Lukuisissa maissa tehdyissä tutkimuksissa on havaittu, että korkeasti koulutettujen ja keskiluokkaisten vanhempien lapset menestyvät alemmin koulutettujen ja työväenluokkaisten vanhempien lapsia paremmin koulutuspoluillaan. (Kalalahti ym., 2015a, 23.)

Kouluvalintamahdollisuuden ja sen käyttämisen vaikuttavat esimerkiksi ottavan erityisen vakavasti varsinkin ne vanhemmat, joilla on keskimääräistä enemmän (koulutus)kulttuurista pääomaa tai koulutustietoutta. Näillä vanhemmilla on muun muassa tietoa erilaista mahdollisuuksista, ja he myös käyttävät oikeutta kouluvalintamahdollisuuteensa ahkerasti. (Silvennoinen, Seppänen, Rinne, & Simola, 2012, 504). Kouluvalintaoikeuttaan voimakkaimmin käyttävät vanhemmat – joiden lapsi on painotetussa opetuksessa muussa kuin lähikoulussa – suhtautuvat muita vanhempia myönteisemmin koulumarkkinoihin ja koulutuksessa tapahtuvaan kilpailuun. Puolestaan ne vanhemmat, joiden lapsi on lähikoulun yleisopetuksen ryhmässä, suhtautuvat kaikista kielteisemmin koulumarkkinoihin ja kannattavat yhtenäistä peruskoulua. Lisäksi vaikuttaa siltä, että valintaoikeuttaan eniten käyttävät vanhemmat eivät pidä lähikoulujensa mahdollisuuksia tarjota tasa-arvoisia mahdollisuuksia kaikille lapsille kovin hyvinä, kun taas vastaavasti lähikoulua ilman erityistä painotusta käyvien lasten vanhemmat näkevät tilanteen toisin. (Kalalahti & Varjo, 2016, 55–58.) Vertailtaessa suomalaisten ja saksalaisten vanhempien näkemyksiä kouluvalinnasta, suomalaisten vanhempien näkemys ”kilpailuyhteiskuntaa” – sisältäen kilpailun kouluista – kohtaan oli saksalaisia kielteisempi. Suomalaiset vanhemmat eivät halua huippukouluja, mutta kannattavat kuitenkin erikoistuneita lähikouluja. (Lobato, Bernelius, & Kosunen, 2018)

Näyttää siltä, että kouluttautuneella keskiluokalla on olemassa ”oikeanlaisia” resursseja sekä sosiaalisen että kulttuurisen pääoman saralla, jotta he esimerkiksi pystyvät tulkitsemaan koulujen lähettämiä viestejä tai selviämään monimutkaisistakin hakuprosesseista, mutta toisaalta myös keräämään ja käyttämään hyödykseen sisäpiiritietoa tai jonkinlaista hiljaista

tietoa kouluista ja niiden eroista. Lisäksi nämä vanhemmat ovat aktiivisia etsimään tietoa erilaisista mahdollisuuksista. (Kosunen, Carrasco, & Flores, 2015; Seppänen, Rinne, & Riipinen, 2012, 229–230.) Vanhemmat, joilla on paljon koulutusresursseja ovat myös tietoisia koulujen välisistä eroista ja tiedostavat kouluvalinnoilla olevan merkitystä. (Lobato ym., 2018.) Yleisesti ottaen vanhemmat, joilla on hyvin taloudellisia, sosiaalisia ja kulttuurisia resursseja eli pääomia, ovat paremmin varustautuneita valitsemaan ja saamaan lapselleen toivotun koulun kuin vähemmän näitä resursseja omaavat vanhemmat (van Zanten, 2015, 22).

Kansainvälisissä tutkimuksissa on havaittu, että työväenluokkaiset vanhemmat eivät välttämättä havaitse koulujen maineiden välistä hierarkiaa samalla tavalla kuin keskiluokkaiset, vaikka eri yläkouluihin siirtyvät lapset olisivatkin hyvin tietoisia koulujen välisistä eroista. (Seppänen ym., 2012, 229–230.) Yhteiskuntaluokkien määrittely voidaan Suomessa tehdä monella eri tavalla, mutta jos käytetään tulkintaa, jonka mukaan Suomessa yli kolmannes ammatissa toimivista henkilöistä kuului vuonna 2015 työväenluokkaan, karkeasti puolet ylempään ja alempaan keskiluokkaan ja noin 15 % yrittäjiin (Melin, 2019), eli kärjistetysti jopa yli kolmannes suomalaisista koululaisten vanhemmista ei havaitsisi kovinkaan hyvin erilaisia koulujen välisiä mainehierarkioita. Erot eivät tietenkään todellisuudessa kulje jyrkästi eri luokkarajojen välissä ja luokatkin ovat hyvin liukuva käsite, mutta silti voidaan ajatella, että suuri osa suomalaisista vanhemmista on heikossa asemassa erilaisten koulujen tunnistamisen suhteen ja tätä kautta alakynnessä kouluvalintamarkkinoilla.

Toisessa päässä puolestaan korkeassa sosioekonomisessa asemassa olevat vanhemmat vaikuttaisivat puolestaan tiedostavan jopa kouluvalinnan vaikutukset koulutustulosten eriytymiseen. Samalla he ovat kuitenkin vähiten innokkaita pitämään peruskoulun kaikille samanlaisena. Nämä vanhemmat haluaisivat lisää mahdollisuuksia erityisesti lahjakkaille lapsille, mutta kannattavat kuitenkin myös mahdollisimman tasalaatuista peruskoulua kaikille, unohtamatta kuitenkaan mahdollisuuksia kilpailla koulutuksessa erilaisten kykyjen ja ominaisuuksien avulla. (Kalalahti ym., 2015b; Koivisto, 2018; Kosunen, Seppänen, & Rinne, 2015). Tulokset korkeassa sosioekonomisessa asemassa olevien vanhempien ajatuksista vaikuttavat keskenään ristiriitaisilta, mutta saattavat kertoa esimerkiksi siitä, että vanhemmat ovat tietoisia kouluvalinnanvapauden huomattavankin erilaisista vaikutuksista yksilön ja toisaalta myös yhteiskunnan tasolla. Tuloksista voitaneen tulkita, että nämä vanhemmat kannattavat tasa-arvoista koulutusta, mutta tiedostavat nykyisen järjestelmän ongelmat ja pyrkivät keräämään järjestelmän puutteista koituvan hyödyn omille jälkeläisilleen mahdollisuuksien näin salliessa, toivoen kuitenkin, että järjestelmän virheet voitaisiin samaan

aikaan korjata. Samansuuntaista tulkintaa ovat tehneet muun muassa Kalalahti ja kumppanit (2015b, 387).

Kouluvalinnalla vaikuttaisi siis olevan suuri merkitys koulujen eriytymisen mahdollistajana. Koulujen eriytyminen tapahtuu muun muassa kouluvalinnan kautta monin erilaisin prosessein. Kalalahden ja kumppanien (2015b, 382) analyysissä tultiin lopputulemaan, jonka mukaan kouluvalinta on jakautunut Suomessa ”*samoin kuin jyrkempien yhteiskuntasidonnaisten luokkien maissa.*” Tämä ei ole miellyttävä ajatus suomalaisen peruskoulun eriytymiskehitystä ajatellen, sillä voidaan pitää hyvin selkeänä sitä, että kouluvalinta lisää myös koulujen eriytymistä. Tästä osoituksena esimerkiksi se, että painotetuille luokille hakeutuvat ja pääsevät useammin keskiluokkaan kuuluvien perheiden lapset. Esille voidaan nostaa ennen kaikkea se, että tiettyihin kouluihin hakeutuu sosioekonomisesti hyvässä asemassa olevien perheiden lapsia. Kun lisäksi tiedetään, että koulutustaso Suomessa periytyy yhä edelleen melko voimakkaasti, ja sitä ohjaavista tekijöistä merkittävimpiä ovat vanhempien kulttuurinen ja sosiaalinen pääoma (esim. Kärkkäinen, 2004, 98–99), näyttää siltä, että sosioekonomiset taustatekijät yhdessä kouluvalinnan kanssa ovat merkittäviä koulujen eriytymiseen johtavia tekijöitä. Perheen sosioekonomisella asemalla on melko selkeä yhteys myös siihen, millaisia oppimistuloksia oppilaat saavuttavat. Suomessakin yhteys on olemassa, vaikkakin se on pienempi verrattuna esimerkiksi naapurimaihimme Ruotsiin ja Viroon (Pöder, Lauri, & Veski, 2017, 683). Seuraavassa luvussa käsitellään koulujen oppilaspohjien sosioekonomista eriytymistä erityisesti alueellisesta näkökulmasta.

2.4 Yhteiskunnallinen polarisaatio, alueellinen eriytyminen ja koulut

Koulut voivat siis eriytyä vanhempien tekemien kouluvalintojen kautta oppilaspohjiltaan. Oppilaspohjalla tarkoitetaan tässä yhteydessä koulun oppilaiden sosioekonomisia taustatekijöitä. Koulun oppilaspohjan muodostumiseen vaikuttaa se, millaisia oppilaita koulun lähialueella asuu, mutta myös se, millaisia taustoja omaavia oppilaita kouluun hakeutuu kauempaa. Tässä luvussa avataan hieman tarkemmin prosessia, jonka perusteella koulujen oppilaspohjat muotoutuvat Suomessa.

Kuten aiemmin todettiin, on havaittu, että perheen yhteiskuntaluokka vaikuttaa vahvasti kouluvalintaprosesseihin Suomessa. Erityisesti ero tulee ilmi lasten hakeutumisessa painotetuille luokille, mutta näkyy myös hakeutumisessa oman oppilasalueen ulkopuoliseen kouluun. Muun muassa koulun koetun paremmuuden toisiin kouluihin verrattuna – koulujen

suosiohierarkian – yhteys perheiden kouluvalintapäätöksiin on osoitettu. Suosittuihin kouluihin hakevat selkeästi enemmän korkeammassa sosioekonomisessa asemassa olevien vanhempien lapset ja ”torjuttujen” koulujen alueilta vastaavat hakeutuvat pois päin. (kts. esim. Leivo, 2008, 37–40; Seppänen ym., 2012.)

Oppilas pohjan sosioekonominen tausta yhdistää kouluvalinnat kaupunginosien sosiaaliseen rakenteeseen (Bernelius, 2013b). Eräässä suomalaistutkimuksessa havaittiin, että eri alueilla sijaitsevien koulujen välillä on merkittäviä eroja oppilaiden sosioekonomisissa taustoissa: Tutkimuksen mukaan sosioekonomisesti matalimman oppilaiden taustaprofiilin omaaviin kouluihin valikoitui oppilaita alueilta, joilla keskimäärin vain 11 %:lla aikuisväestöstä oli korkeakoulututkinto. Vastaavasti korkeimman sosioekonomisen profiilin kouluihin oppilaita puolestaan hakeutui alueilta, joiden aikuisväestöllä 36 %:lla oli korkeakoulututkinto. Määrä on yli kolminkertainen ja osoittaa, kuinka kouluissa työskennellään eriytyneissä toimintaympäristöissä. (Kosunen, Bernelius, Seppänen, & Porkka, 2018, 15–16.) Vaikka kouluille ei olisikaan asetettu erityisiä oppilasalueita, muodostuu koulujen ensisijainen oppilas pohja yleensä sijaintialueen väestön perusteella (Bernelius, 2013a, 120; 2013b). Yhteiskunnallisessa keskustelussa onkin jo ainakin vuosituhanen alkupuolelta alkaen nähty ongelmalliseksi se, että koulutettujen ja hyvin toimeentulevien lapset kasautuvat tiettyihin kouluihin, sosioekonomisesti heikommassa olevien vanhempien lapset puolestaan toisiin kouluihin. (Seppänen, 2006, 182.; Seppänen ym., 2012, 229.)

Koulujen naapurustojen sosioekonominen ja etninen eriytyminen vaikuttavat koulun suosioon ja lähtökohtaisesti niin sanotuilla hyväosaisilla alueilla sijaitsevien koulujen on helpompi houkutella oppilaita myös muilta alueilta, kun taas huonompiosaisilla alueilla sijaitsevilla kouluilla on vaikeampaa tältä osin (Bernelius, 2013a, 120; 2013b). Suurimmat erot perheiden välillä näkyvät äitien tulojen kautta. Esimerkiksi Seppänen ja kumppanit ovat todenneet tutkimuksessaan, että: *”muuhun kuin asuinalueensa torjuttuun yläkouluun haki hyvä- ja suurituloisten äitien lapsista lähes kaksi kolmesta (63 %), pieni- ja keskituloisten äitien lapsista vain reilu kolmannes (35 %). Vastaavasti torjuttujen koulujen alueilla asuvista korkeakoulutettujen äitien lapsissa muuhun kouluun hakijoita oli reilu puolet (55 %) ja heitä matalammin koulutettujen äitien lapsista vajaa kolmannes.”* (Seppänen ym., 2012, 237.)

Bernelius (Rossi, 2018) väittää, että keskiluokkaisten ihmisten valinnat ohjaavat koulujen ja alueiden kehittymistä. On havaittu, että tietyiltä alueilta muuttavat pois ne, joilla on siihen mahdollisuus esimerkiksi juuri taloudellisten resurssien puolesta. Samaan aikaan alueelle

jäävät ”jumiin” ne – tyypillisesti huono-osaisemmat – joilla ei ole mahdollisuutta syystä tai toisesta muuttaa pois (Nevalainen, 2018, 4). Varsinaisesta ”white flight”-ilmiöstä, eli valkoisten, hyvin toimeentulevien perheiden muuttamisesta paremmille alueille ei kuitenkaan Suomessa voida puhua, vaikka jonkinlaista niin sanottujen huonojen alueiden välttelyä esiintyykin (Komulainen, 2012, 60).

Suomalaisten perheiden asuinpäätöksiin ei vaikuta olevan kovin suurta yhteyttä sillä, asuuko alueella esimerkiksi paljon maahanmuuttajia, eikä koulukaan nouse kovin merkittäväksi tekijäksi ainakaan alakouluikäisten lasten perheiden muuttopäätösten kohdalla. (Komulainen, 2012, 55–56.) Asuinaluetta valitessa vanhempien päätökseen näyttäisivät vaikuttavan ensin muut tekijät, kuten asunnon ominaisuudet, alueen turvallisuus ja rauhallisuus, asumiskustannukset ja työ- sekä koulumatkojen pituus (Dhalmann, Vaattovaara, & Vilkkama, 2013.) Koulun hyvä taso ei noussut Dhalmannin ja kumppaneiden (2013) tutkimuksessa tärkeään osaan, sillä sen nimesi tärkeimmäksi syyksi vain noin yksi prosentti tutkimukseen osallistuneista vanhemmista, neljän tärkeimmän kriteerin joukkoon se nousi vain noin yhdeksällä prosentilla vastaajista.

Koulujen välisillä eroilla on kuitenkin merkitystä, ja erot siirtyvät asuntomarkkinoiden kautta jopa asuinalueiden yleiseen hintakehitykseen. Helsingissä koulujen väliset erot heijastuvat asuntojen hintoihin vastaavalla suuruudella, kuin esimerkiksi Isossa-Britanniassa tai Yhdysvalloissa. Näissä maissa koulujen väliset laatuerot ovat kuitenkin paljon suurempia kuin Suomessa. Karkeasti asuntojen hintakehityksen yhteys koulujen laatuun on sellainen, että yhden keskihajonnan nousu standardoidussa koepistemäärässä vaikuttaa asuntojen hintoihin korottavasti noin 2,5 prosenttiyksikköä. (Saarimaa, Harjunen, & Kortelainen, 2014.) Samaisen tutkimuksen tulokset osoittivat myös, että vanhemmat havaitsevat selkeitä eroja alakoulujen välillä, vaikka erot oppilaiden tuloksissa ovatkin Suomessa matalia kansainvälisessä tarkastelussa. Muutto- tai asunnonostopäätöksiin liittyvät tekijät koulujen suhteen vaikuttaisivat kuitenkin olevan ennen kaikkea koulussa opiskelevien lasten sosioekonomiseen taustaan liittyviä, sillä vanhempien on vaikea saada tietoa koulujen oppimistuloksista. (Saarimaa ym., 2014, 1, 3, 20–22.) Tästä seuraa, että vanhemmat muodostavat mielikuvansa koulusta muilla tavoin kuin oppimistuloksiin pohjautuen.

Asuntomarkkinoiden rooli alueellisen segregaatoin ja tätä kautta myös koulujen lähialueiden sosioekonomisen pohjan eriytymisen kohdalla on melko suoraviivainen: hyvin toimeentulevilla on yksinkertaisesti paremmat mahdollisuudet valita asuinpaikkansa, sillä

heillä on yleensä paremmin varaa muuttaa erilaisille alueille. Vaikka yhteiskunnallinen turvaverkko Suomessa oletettavasti tasoittaakin tuloeroista johtuvaa alueellista eriytymistä, ovat jotkin sosiaaliryhmät toisia paremmassa tai heikommassa asemassa asuntomarkkinoilla.

Siinä missä keskiluokkaisten ihmisten valintamahdollisuudet ovat yleisesti hyviä, ovat esimerkiksi maahanmuuttajat usein kantasuomalaisia heikommassa asemassa. Eduskunnan tarkastusvaliokunnan 14.2.2018 saamassa lausunnossa (Nevalainen, 2018) käsitellään alueellista eriytymistä ja segregaatiota maahanmuuttajien asuntomarkkinoiden näkökulmasta. Lausunnossa nostetaan esille segregaation syitä ja seurauksia. Kaupunkien sisäisen sosioekonomisen segregaation voidaan katsoa olevan lausunnon mukaan seurausta tulonjaon, työllisyyden, koulutuksen ja hyvinvoinnin epätasaisesta jakautumisesta. Nämä erot siirtyvät ja keskittyvät asuntomarkkinoiden vaikutuksesta kaupunkien eri alueille. Kallis asuminen heikentää heikommassa asemassa olevien, kuten työttömien ja maahanmuuttajien mahdollisuuksia valita asuinpaikkansa, jolloin asuntomarkkina ”työntää” heidät heikoille alueille tai vaikeuttaa alueelta pois lähtemistä. Lausunnon mukaan osattomuus hyvinvoinnista, sisältäen muun muassa koulutuksen, terveyden, tulot ja työllisyyden, keskittyy usein asuntomarkkina-asemaltaan heikoille alueille. (Nevalainen, 2018, 2.) Joissakin eurooppalaisissa tutkimuksissa on saatu selville, että huono-osaisille alueille syntyvät lapset viettävät siellä yleensä koko lapsuutensa ja myös jäävät sinne aikuistuttuaan. (Nevalainen, 2018, 4).

Uusimpana käänteenä eriarvoistumiskehitystä vauhdittanee koronaviruspandemia. Terveysten ja hyvinvoinnin laitoksen tutkimuspäällikkönä toimiva Mari Vaalavuo nosti sekä Helsingin Sanomien mielipidepalstalla, että THL:n verkkosivujen blogissa julkaistussa kirjoituksessaan esille koronaviruksen esiintuoman ja osaltaan myös tuottaman eriarvoisuuden kasvun. Vaalavuo toteaa muun muassa, että alempiin sosioekonomisiin ryhmiin kuuluvilla ihmisillä on usein enemmän perussairauksia kuin ylempiin luokkiin kuuluvilla, minkä vuoksi he ovat suuremmassa riskiryhmässä viruksen suhteen. Lisäksi näihin ryhmiin kuuluu keskimääräistä enemmän sellaisia työntekijöitä, joilla ei työnsä luonteen vuoksi ole mahdollisuutta välttää lähikontakteja tai jäädä etätöihin. Näissä töissä toimivilla on suurempia vaikeuksia muokata työtehtäviään kriisin tuottamien rajoitusten tai oman tilanteensa mukaan, minkä lisäksi alempiin sosioekonomisiin luokkiin kuuluvien työnantaja ei välttämättä ole yhtä joustava kuin korkeampiin luokkiin kuuluvilla voi olla. Pienituloisilla ongelmia voi lisäksi aiheuttaa pienestä palkasta johtuva säästöjen kerryttämisen vaikeus. (Vaalavuo, 2020a; Vaalavuo, 2020b.)

Jo pari vuotta ennen koronaviruksen aiheuttamaa erityistilannetta Sipilän hallituksen nimeämä eriarvoisuutta käsittelevä työryhmä totesi raportissaan, että eriarvoistumiskehitys Suomessa on ollut haasteellista 2010-luvun jälkimmäisellä puoliskolla, sillä valtiovallan mahdollisuudet tukea heikossa asemassa olevia kansalaisiaan ovat huonontuneet. (Valtioneuvoston kanslia 2018, 12). Työryhmä mainitsi raportissaan, että yhteiskunnan eriarvoistumisen kehityksen osoittavia tekijöitä ovat muun muassa toimeentulotukiasiakkuuden piteneminen, eri väestöryhmien väliset terveyserojen kasvu, äänestämisen ja politiikassa mukana olemisen erot eri ryhmien välillä, korkea työttömyys, köyhyys, koulutuksen ja työelämän ulkopuolella olevien nuorten merkittävä osuus, ylivelkaantuminen sekä maahanmuuttajien kotouttamisen haasteet. Lisäksi aiemmissa tutkimuksissa ja selvityksissä on havaittu muun muassa toimeentulotuen tarpeen kasvua, sekä tuloerojen kasvua Sipilän hallituksen työryhmän raporttia edeltäneinä vuosina 2015–2017 (Mukkila, Ilmakunnas, Moisio, & Saikkonen, 2017).

Näyttää siltä, että työttömien, lapsiperheiden ja opiskelijoiden toimeentulo on huonontunut jo vuosien 2012–2018 välisenä aikana (Honkanen, 2018). Myös Helsingin yliopiston kaupunkimaantieteen apulaisprofessori Venla Bernelius kirjoitti Alue- ja ympäristötutkimuksen seuran, Yhteiskuntatieteellisen ympäristötutkimuksen seuran ja Suomen maantieteellisen seuran yhdessä kustantamassa Versus-verkkojulkaisussa (2019), että hyvinvoinnin yhteiskunnallisen polarisaation tuottamat haasteet Suomessa ovat lisääntyneet, ja että *“kaupunkinaapurustojen eriytyminen herättää aiheellisesti huolta myös koulujen oppimisympäristöjen ja oppimistulosten eriytymisen näkökulmasta.”* Ilmiö ei ole uusi: hyvämaineisten ja laadukkaiksi nähtyjen koulujen alueilla yliedustettuina olivat jo vuosituhannen ensimmäisellä vuosikymmenellä korkeasti koulutettujen vanhempien perheet (Seppänen, 2006, 258–260).

Yhteiskunnallista polarisaatiota ja eriarvoistumista voidaan tutkia monesta eri näkökulmasta ja usein eri mittarein. Valtioneuvoston eriarvoisuutta käsittelevä työryhmä on viitannut eriarvoisuudella kolmeen toisiinsa kytkeytyvään asiakokonaisuuteen; 1. ihmisten pärjäämiseen yhteiskunnassa, 2. ihmisten kykyyn käyttää elinkeinoelämän ja julkisen vallan tarjoamia mahdollisuuksia, 3. sosiaalisiin tilanteisiin, joissa henkinen autonomia ja fyysinen terveys jäävät toteutumatta. Esimerkiksi sosiaalisen liikkuvuuden toteutumiseen vaikuttavat tekijät, byrokratia ja kannustinloukut sekä pitkäaikainen köyhyys ja lyhyeksi jäävä elinajanodote ovat esimerkkejä edellä mainittujen asiakokonaisuuksien sisältämistä asioista. (Valtioneuvoston kanslia 2018, 13.)

Tasa-arvo ja eriarvoisuuskeskustelu rajautuu Saaren (2019, 22) mukaan yksinkertaisimmillaan pohdintaan mahdollisuuksien ja tulosten tasa-arvon välillä. Selvää kuitenkin on, että koulujen voimakas eriytyminen ei ole toivottavaa yhteiskuntamme tasa-arvoisuuden kannalta. Perusteluna voidaan käyttää vaikkapa sitä, että korkealle kouluttautuminen ja hyvä menestys koulutuksessa näkyvät myöhemmin hyvinä mahdollisuuksina työelämässä ja siirtyvät sitä kautta osaksi koko elämäntiekkoa. Edelliseen viitaten suomalaisessa työelämäkulttuurissa koulutusta on perinteisesti arvostettu korkealle. Koulutuksen arvostus näkyy monella tavalla, sillä koulutustason noustessa työllistymismahdollisuus nousee, samoin kuin vastaavasti työttömyyden ja työmarkkinoiden ulkopuolelle putoamisen riski laskee. Tämä pätee kaikissa ikäluokissa, mikä kertoo siitä, että kyseessä ei ole uusi ilmiö. Kehityskulku näyttäisi tuoreimpienkin tilastojenkin valossa jatkavan samalla linjalla, eli mitä korkeampi koulutus, sitä varmemmalla todennäköisyydellä töihin pääsee. (kts. esim. Kalenius, 2014; Suomen virallinen tilasto, 2018.) Lisäksi Tilastokeskuksen tiedot osoittavat, että Suomessa pidemmälle kouluttautuminen ennustaa parempia tuloja, ja maisteriksi valmistuneet tienaavatkin keskimäärin paremmin kuin saman ikäiset matalammin koulutetut (kts. esim. Kivinen, Hedman, & Kaipainen, 2012, 564; Suomen virallinen tilasto, 2016; 2019; 2020).

Yhteiskunnallinen polarisaatio ja alueellinen eriarvoistuminen liittyvät voimakkaasti toisiinsa. Koulujen eriytymiseen yhteiskunnallinen polarisaatio puolestaan linkittyy ehkä voimakkaimmin juuri alueellisen eriarvoistumisen kautta. Hyvin toimeentulevilla ja pärjäävillä perheillä on usein hyvät mahdollisuudet valita asuinpaikkansa ja tätä kautta saada lapsensa parempien koulutusmahdollisuuksien äärelle. Näillä koulutusmahdollisuuksilla voidaan tarkoittaa esimerkiksi painotettua opetusta tai suurempia valinnaisaineiden määriä tietyissä kouluissa. Vastaavasti heikossa asemassa olevat voivat usein joutua tosiasiaa tyytymään heikompiin vaihtoehtoihin lasten koulutuksen suhteen, vaikka periaatteessa valintamahdollisuudet olisivatkin samat. Toisaalta on hyvin haastavaa aukottomasti todeta, johtuuko hyvien koulujen ja huonojen koulujen olemassaolo siitä, että aktiiviset, lastensa koulutuksesta kiinnostuneet vanhemmat luovat tietyille alueille toiminnallaan hyvämaineisen koulun, vai houkutteleeko hyvämaineinen ja hyväksi koettu koulu näitä vanhempia muuttamaan alueelle lapsineen. Joka tapauksessa vaikuttaa siltä, että yhteiskunnallinen polarisaatio heijastuu Suomessa kouluihin pitkälti aluevaikutusten kautta. Erillinen kysymys on kuitenkin se, kuinka suuria rooli kouluilla on alueellisen segregaaion syntymisessä, vai muodostuuko koulujen eriytyminen alueellisen segregaaion kautta. Todennäköisesti vaikutus on kumpaankin suuntaan.

3 Aineisto ja metodologia

Tutkielmassa toteutettava tutkimus on luonteeltaan empiirinen eli havainnoiva. Empiirisessä tutkimuksessa testataan, toteutuuko tutkimusteorioiden pohjalta muodostettu hypoteesi eli oletamus käytännössä, tai siinä voi olla tavoitteena pyrkiä selvittämään jonkin ilmiön syitä tai pyrkiä löytämään ratkaisuja jonkin asian toteuttamiseksi (Heikkilä, 2014, 12). Empiirinen tutkimus pyrkii yksittäistapausten pohjalta löytämään jonkinlaisia säännönmukaisuuksia tai lainalaisuuksia aineistosta (Valli, 2015). Empiirinen tutkimus kuitenkin pohjaa aikaisempaan tutkimustietoon ja tutkijan omaan aivotoimintaan teoretisoinnissa ja mittaustulosten analyysissä. (Jokivuori & Hietala, 2014).

Tässä pro gradu -tutkielmassa tutkimusote on kvantitatiivinen eli määrällinen tai tilastollinen. Tilastollinen tutkimus on yksinkertaistettuna lähinnä numeroiden hyödyntämistä ja hyväksikäyttöä tutkimuksessa. Tutkimusaineistoa käsitellään numeroin erilaisten matemaattisten toimenpiteiden kautta hyödyntäen esimerkiksi erilaisia tilasto-ohjelmia. Tilastollinen tutkimus perustuukin nimensä mukaisesti tilastotieteen käyttöön. Tilastotiede puolestaan on empiirisen eli kokemuseräisen ja numeerisen tietoaineiston hankinnan suunnittelua, tietojen keräämistä, esittämistä sekä analysointia käsittelevää tiedettä. Tilastotiede takaa tulosten tarkastelulle luotettavan pohjan, sekä mahdollistaa tiedon tiivistämisen ja kuvailemisen, sekä erilaisia menetelmiä päättelyyn ja matemaattiseen mallintamiseen. (Valli, 2015.) Kvantitatiivisessa eli tilastollisessa tutkimuksessa pyritään selvittämään esimerkiksi lukumääriin, prosenttiosuuksiin tai asioiden tilastollisiin riippuvuussuhteisiin liittyviä asioita. Kvantitatiivisessa tutkimuksessa pyritään usein myös yleistämään aineistosta saatuja tuloksia koskemaan aineistoa suurempaa joukkoa. Tällöin kvantitatiivinen tutkimus voi olla luonteeltaan myös ennustavaa. Kvantitatiivisen tutkimuksen heikkoutena voidaan nähdä se, että sen avulla ei yleensä pystytä selittämään asioiden syitä, vaan ainoastaan toteamaan voimassa oleva tilanne ja luomaan malleja, joiden avulla voidaan ennustaa tutkittuja havaintoyksiköitä laajemmassa joukkiossa esiintyviä asioita. (Heikkilä, 2014, 14–15.) Tulosten tulkinnat puolestaan syntyvät nimenomaan *tulkinnan* kautta. Tutkijan onkin tärkeätä kyetä pitämään havaitut tulokset ja niistä tehdyt tulkinnat ja johtopäätökset erillään toisistaan. (Jokivuori & Hietala, 2014.)

Tässä tutkielmassa pyritään ennen kaikkea selvittämään syitä siihen, minkä vuoksi koulut vaikuttaisivat eriytyvän Suomessa myös alueellisesti ja selittävätkö samat jo aiemmissa tutkimuksissa havaitut tekijät kehitystä samalla tavoin kaikkialla Suomessa, vai onko kyseessä

tiettyjen alueiden erityispiirre. Tutkielman on osittain myös kausaalinen eli selittävä, sillä se pyrkii ymmärtämään ilmiöiden välisiä syy- ja seuraussuhteita eri muuttujien välisten riippuvuuksien avulla (Heikkilä, 2014, 14). Yhteiskunta- tai ihmistieteellisessä lähestymistavassa ei olla sinänsä kiinnostuneita tunnusluvuista ja numeroista, vaan kiinnostus kohdistuu syihin niiden takana. Tutkimuksessa ollaan siis kiinnostuneita ilmiöiden välisistä yhteyksistä, joita pyritään todentamaan kvantitatiivisten menetelmien avulla. Tämä on hyvin tyypillinen lähestymistapa nimenomaan empiirisissä yhteiskunta- ja ihmistieteellisissä tutkimuksissa. (Jokivuori & Hietala, 2014.) Kyseessä on poikkileikkaustutkimus, sillä se tutkii koulujen oppimistulosten ja tiettyjen sosioekonomisten tekijöiden yhteyttä yhdessä hetkessä, eikä pyri pitkittäistutkimuksen tapaan seuraamaan muuttujien yhteyden kehittymistä tiettyä ajanjaksona. (Heikkilä, 2014, 14.)

Tieteellisessä tutkimuksessa ei pitäisi milloinkaan pyrkiä tarttumaan ensimmäiseen tarjolla olevaan selitykseen. Hyvässä tutkimuksessa tulisi pyrkiä siihen, että todella selvitetään tutkittavaan ilmiöön liittyvien tekijöiden välisiä yhteyksiä ja eroja. Usein tutkimusongelmat ovat kuitenkin niin monitahoisia, että on tarkasteltava yhtä aikaa useita muuttujia ja selittäviä tekijöitä. (Jokivuori & Hietala, 2014.) Tässä tutkimuksessa mahdolliset selittävät tekijät ovat hyvin moninaisia ja liittyvät usein hyvin vahvasti toisiinsa. Tämän vuoksi kvantitatiivisten menetelmien kautta saatujen tulosten tulkinnassa on tärkeää tuntee hyvin tutkimusongelman taustateoria ja aiempien tutkimusten tuloksia. Tämä auttaa, kun tehdään saaduista tuloksista johtopäätöksiä ja pyritään yhdistämään asioita reaali maailmaan.

Monipuolisella eri tieteenalojen yhteistyöllä on mahdollista tulkita tuloksia eri näkökulmista ja pyrkiä löytämään kaikista vakuuttavimmat selitykset. Tutkimuskentällä aiemmat tutkimukset koulujen eriytymisestä asemoituvat usein kaupunkimaantieteen, kasvatustieteen ja (kasvatus)sosiologian välimaastoon, ollen näin monitieteisiä sekä metodeiltaan ja monipuolisilta lähestymiskulmiltaan. Myös tässä tutkielmassa hyödynnetään monitieteisiä metodeja. Tausta-aineistona ja teoriapohjan luomisen perustana käytetään erityisesti aiempaa suomalaista tutkimusta, sillä tutkimuksen konteksti on kuitenkin suomalaisissa kaupungeissa tapahtuva koulujen oppimistulosten eriytyminen ja sen yhteys alueellisiin sosioekonomisiin muuttujiin. Kansainvälisten tulosten rooli on toimia lähinnä vertailukohtana muodostettaessa kuvaa koulujen eriytymisen ilmiöstä Suomessa, selityksiä näistä tutkimuksista ei kuitenkaan pyritä juurikaan hakemaan.

3.1 Aineiston kuvaus

Tutkimusaineistona tutkimuksessa toimivat Suomen tärkeimmän tilastojen tuottajan Tilastokeskuksen (Heikkilä, 2014, 16) tuottamat yhteiskuntarakenteen seurannan aineistot (YKR-aineistot), sekä Oulun kaupungin sivistys- ja kulttuuritoimen toimittamat Oulun yläkoulujen päättökeskiarvot ilmoittava numeroaineisto. Päättökeskiarvot saatiin tutkimuskäyttöön hakemalla Oulun kaupungin sivistys- ja kulttuuripalveluilta tutkimuslupaa kirjallisesti. Tutkimuslupa myönnettiin 27.05.2020 hakemuksen mukaisesti, jonka jälkeen kaupunki toimitti yläkoulujen päättökeskiarvot tutkimuskäyttöön.

Päättökeskiarvoja päädyttiin käyttämään tutkimuksessa oppimistuloksia kuvaavana muuttujana, sillä standardoituja testejä, joiden avulla oppimista olisi voitu arvioida, ei Oulussa ole toteutettu. Päättökeskiarvoa on lisäksi käytetty aiemminkin mittaamaan lapsen koulumenestystä Suomessa (esim. Kalalahti ym., 2015a). Lisäksi tutkielmassa käytettävien aineistojen joukkoon lukeutuvat 7.-9.-luokan opetusta antavien koulujen sijaintitiedot ja Oulun kaupungin niin sanottujen ”SIKU” eli sivistys- ja kulttuurialueiden rajat.

YKR-aineistosta käytettäviä tietoja ovat 250x250 metriä kuvaaviin tilastoruutuihin yhdistetyt tiedot alueen asukkaiden rahatulosta sekä työssäkäyvien koulutustasoista. Tilastoruudut esitetään paikkatietoihin sidottuina Suomen kartalla. Tutkimuksessa kiinnostuksen kohteena on kuitenkin vain Oulun kaupunki, joten kaupungin rajojen ulkopuolelle jäävät ruudut rajattiin pois hyödyntäen Oulun kaupungin hallintoalueen rajoja, jotka poimittiin Maanmittauslaitoksen avoimen datan aineistoista shapefile-tiedostona.

Käytössä olleen YKR-aineiston tiedot olivat vuodelta 2017, sisältäen muun muassa tutkimuksessa käytettävät tiedot kunkin ruudun asukkaiden koulutusasteista ja ruudun asukkaiden yhteenlasketusta rahatulosta. Aineisto on vuodelta 2017, sillä tuoreempaa aineistoa ei tutkielmaa varten onnistuttu saamaan käyttöön. Näin ollen tutkielmaa toteuttaessa jouduttiin tekemään oletus, että Oulun kaupungin sisäiset sosioekonomiset rakenteet eivät ole vuodesta 2017 vuoteen 2020 mennessä muuttuneet niin radikaalisti, että vuoden 2017 YKR-aineiston käyttäminen yhdessä lukuvuoden 2019-2020 koulujen lähialue-estimaattien ja oppimistulosten kanssa heikentäisi merkittävästi tutkielman luotettavuutta. Tiedossa on kuitenkin, että Oulussa on muun muassa rakennettu vuodesta 2017 vuoteen 2020 mennessä uusia asuinalueita, joten täysin luotettavana 2017 vuoden YKR-aineistoa ei voida missään tapauksessa pitää.

Lisäksi tutkimuksessa käytetään Oulun kaupungin avoimen datan portaalin kautta sähköpostitse kysymällä saatua SIKU-alueiden rajat sisältävää shapefile-tiedostoa. SIKU-

alueiden rajoja käytetään tutkielmassa apuna luotaessa erilaisia malleja koulujen lähialue-estimaateiksi. Edellä mainittujen aineistojen lisäksi tutkielmassa käytetään Maanmittauslaitoksen ylläpitämän Paikkatietoikkunan ja siihen liittyvän Paikkatietohakemiston kautta ladattua koulujen sijainnit koordinaatistossa ilmoittavaa tiedostoa, jonka avulla koulut sijoitettiin paikkatieto-ohjelmassa oikeille paikoille.

Oulun kaupungin yläkoulujen lisäksi mukaan tutkielmassa käytettävien lähialue-estimaattien muodostamiseen päätettiin ottaa myös valtion ylläpitämä Oulun normaalikoulu. Yliopiston harjoittelukouluna toimivalla normaalikoululla on oma oppilasalueensa kaupungin koulujen keskellä. Normaalikoulun rehtori(t) osallistuvat myös kaupungin keskisen ja pohjoisen alueen oppilaiden sijoittamiseen. (Oulun kaupunki, 2018.) Tämän vuoksi tutkielman luotettavuuden voidaan katsoa paranevan, mikäli myös normaalikoululle mallinnetaan oma oppilasalue kaupungin koulujen tapaan. Normaalikoulua ei kuitenkaan käsitellä tutkimuksessa, vaan sen roolina on toimia pisteenä, jonka avulla voidaan rajata lähistöllä sijaitsevien kaupungin koulujen lähialue-estimaattien kokoja vastaamaan paremmin todellisuutta. Tutkielmassa ei siis käsitellä Oulun normaalikoulun oppimistuloksien yhteyksiä alueellisiin muuttujiin kuten kaupungin ylläpitämien koulujen kohdalla, vaan sille mallinnettu oppilasalue jätetään kokonaan käsittelemättä tässä tarkoituksessa.

Oulussa toimii myös joitakin yksityisiä kouluja. Näitä ovat: Oulun kristillinen koulu, Oulun steinerkoulu, sekä ruotsinkielinen Svenska Privatskolan i Uleåborg. Yksityiset koulut muokkaavat osaltaan kaupungin koulujen oppilaspohjia, sillä niihin hakeutuu oppilaita eri alueilta, jolloin nämä oppilaat ovat pois kyseisten alueiden lähikoulujen oppilastaseista. Yksityisiä kouluja ei kuitenkaan otettu mukaan lähialue-estimaattien rakentamiseen, sillä näiden koulujen oppilaspohja muodostuu ympäri Oulua tulevista oppilaista, eivätkä näin ollen kouluille muodostetut lähialue-estimaatit kuvaisi sitä, millaisista ympäristöistä koulun oppilaat pääsääntöisesti tulevat. Lisäksi kaupungin toimittama tilastoaineisto ei luonnollisesti sisällä tietoa yksityisten koulujen oppimistuloksista ja näiden tietojen hankkiminen olisi aiheuttanut tutkielman toteuttamiselle liian paljon lisätyötä saatuun hyötyyn suhteutettuna.

Lähialue-estimaatteja rakentaessa aineistosta päätettiin jättää pois kaupungin omista kouluista Oulun kansainvälinen koulu, sekä Tiernan koulu. Kansainvälinen koulu valitsee oppilaansa hakemuksen ja valintakokeen perusteella, sillä kansainvälisessä koulussa opiskellakseen oppilas tarvitsee riittävän englannin kielen taidon. Tästä syystä koulun oppilaat tulevat ympäri Oulua tai joissain tapauksissa jopa lähikunnista, mikäli Oulun kaupunki ja oppilaan kotikunta

tekevät asiasta sopimuksen. (Oulun kaupungin sivistys- ja kulttuurilautakunta, 2019; Oulun kaupunki, 2019, 12, 25). Näin ollen kansainvälisen koulun oppilasaines ei vastaa sitä todellisuutta, mitä voisi olettaa mahdollisen kansainväliselle koululle muodostetun lähialue-estimaatin perusteella. Tiernan koulu puolestaan on kaupungin erityiskoulu, jonka oppilasalueena toimii koko kaupunki, ja sillä on 11 eri toimipistettä ympäri Oulua. (Oulun kaupunki, 2018). Tästä syystä Tiernan koulun mukaan ottaminen ei olisi alueellisen tutkimuksen näkökulmasta kovin hedelmällistä, sillä erilaisilla alueilla sijaitsevat toimipisteet saavat väistämättä erilaisia arvoja lähialue-estimaattiensa mukaisesti, mutta erityiskoulun oppimistulokset on kuitenkin aineistossa ilmoitettu kaikkien toimipisteiden yhteisenä päättökeskiarvona. Näin ollen Tiernan koulun oppimistulosten yhteyden vertailu alueellisiin muuttujiin ei onnistu, vaan jotta vertailua olisi mahdollista tehdä, tulisi tietää koulun eri yksiköiden tulokset erillisinä. Tämä ei tietenkään esimerkiksi oppilaiden tietosuojaan takia ole mahdollista, sillä oletettavasti erityiskoulun oppilasmäärät ovat sen verran pieniä, että joissain tapauksissa jopa oppimistulosten yhdistäminen jopa yksittäisiin oppilaisiin olisi jollain tasolla mahdollista.

Aineistoon valikoitui mukaan 22 koulua ja niitä ympäröivää lähialue-estimaattia. Lähialue-estimaatit muodostettiin kahdella tavalla, jolloin lähialue-estimaattien hienoisen muuttumisen myötä myös niiden sisältä kerätty aineisto muuttui hieman. Seuraavana on esitetty taulukoiden avulla näiden kahden skenaarion pohjalta lähialue-estimaatteihin poimitut tiedot ja niiden tunnusluvut. Taulukossa ”OULU” on esitetty tiedot tilanteessa, jossa lähialue-estimaattien muodostumista ei ole rajattu sivistys- ja kulttuuri-, eli oppilasalueiden rajoja hyödyntäen. ”SIKU” -taulukossa puolestaan lähialue-estimaatteja on rajattu. Tutkimuseettisistä syistä päättökeskiarvojen minimi- ja maksimi-arvot sekä vaihteluväli on jätetty ilmoittamatta.

Taulukko 1. Lähialueiden muuttujien tunnusluvut ilman oppilasaluerajausta.

OULU

	N	Rang e	Mini mum	Maxi mum	Sum	Mean		Std. Deviat ion	Variance
	Statistic	Statis tic	Statis tic	Statis tic	Statisti c	Statist ic	Std. Error	Statisti c	Statistic
PäättöKA	22	-	-	-	180,38	8,199 1	,04679	,21948	,048
rahatulo/ henk	22	8980, 44	1395 5,64	2293 6,09	406444 ,84	18474 ,7655	445,604 00	2090,0 6802	4368384 ,325
rahatulo/a k	22	3223 7,11	2386 8,66	5610 5,77	901756 ,19	40988 ,9178	1787,74 316	8385,2 5869	7031256 3,316
tyol_per/t yol	22	7,16 %	3,51 %	10,67 %	140,84 %	6,4019 %	0,35125 %	1,64752 %	2,714
tyol_kes/t yol	22	43,13 %	28,87 %	72,00 %	1006,65 %	45,756 7 %	2,14936 %	10,0814 1 %	101,635
tyol_kor/t yol	22	49,79 %	17,33 %	67,12 %	1052,51 %	47,841 5 %	2,45568 %	11,5181 6 %	132,668
Valid N (listwise)	22								

Taulukko 2. Lähialueiden muuttujien tunnusluvut oppilasaluerajauksen kanssa.

SIKU

	N	Rang e	Mini mum	Maxi mum	Sum	Mean		Std. Deviat ion	Varia nce
	Stati stic	Stati stic	Stati stic	Stati stic	Statis tic	Statisti c	Std. Error	Statisti c	Statist ic
PäättöKA	22	-	-	-	180,3 8	8,1991	,04679	,21948	,048
rahatulo/henk	22	8523 ,86	1395 5,64	2247 9,50	40575 4,92	18443, 4055	436,818 57	2048,8 6068	41978 30,09 4
rahatulo/ak	22	3223 7,11	2386 8,66	5610 5,77	90190 9,73	40995, 8969	1789,65 508	8394,2 2640	70463 036,7 74
tyol_per/tyol	22	7,08 %	3,59 %	10,67 %	141,63 %	6,4376 %	0,34627 %	1,6241 4 %	2,638
tyol_kes/tyol	22	43,42 %	28,58 %	72,00 %	1009,4 6 %	45,8848 %	2,11877 %	9,9379 3 %	98,76 2
tyol_kor/tyol	22	50,01 %	17,33 %	67,34 %	1048,9 1 %	47,6776 %	2,42399 %	11,369 53 %	129,2 66
Valid N (listwise)	22								

3.2 Aineiston käsittely

Tutkimuksessa käytetään aineistoja, joiden sisältämät tiedot ovat keränneet eri toimijat erilaisiin tarkoituksiin ja erilaisin menetelmin. Onkin tärkeää huomioida, että usein valmiit aineistot eivät ole sellaisessa muodossa, että niitä voisi suoraan käyttää kvantitatiivisissa tutkimuksissa sellaisinaan. Tyypillisesti aineistoja voi joutua muokkaamaan, yhdistelemään ja tarkistamaan ennen niiden käyttämistä tutkimuksessa. (Heikkilä, 2014, 16.)

Tässäkään tutkimuksessa käytettävissä oleva aineisto ei tarjonnut suoraan mahdollisuuksia vertailla erilaisten alueiden eroja toisiinsa. Tämä johtui siitä, että tutkimuksella on vahva yhteys

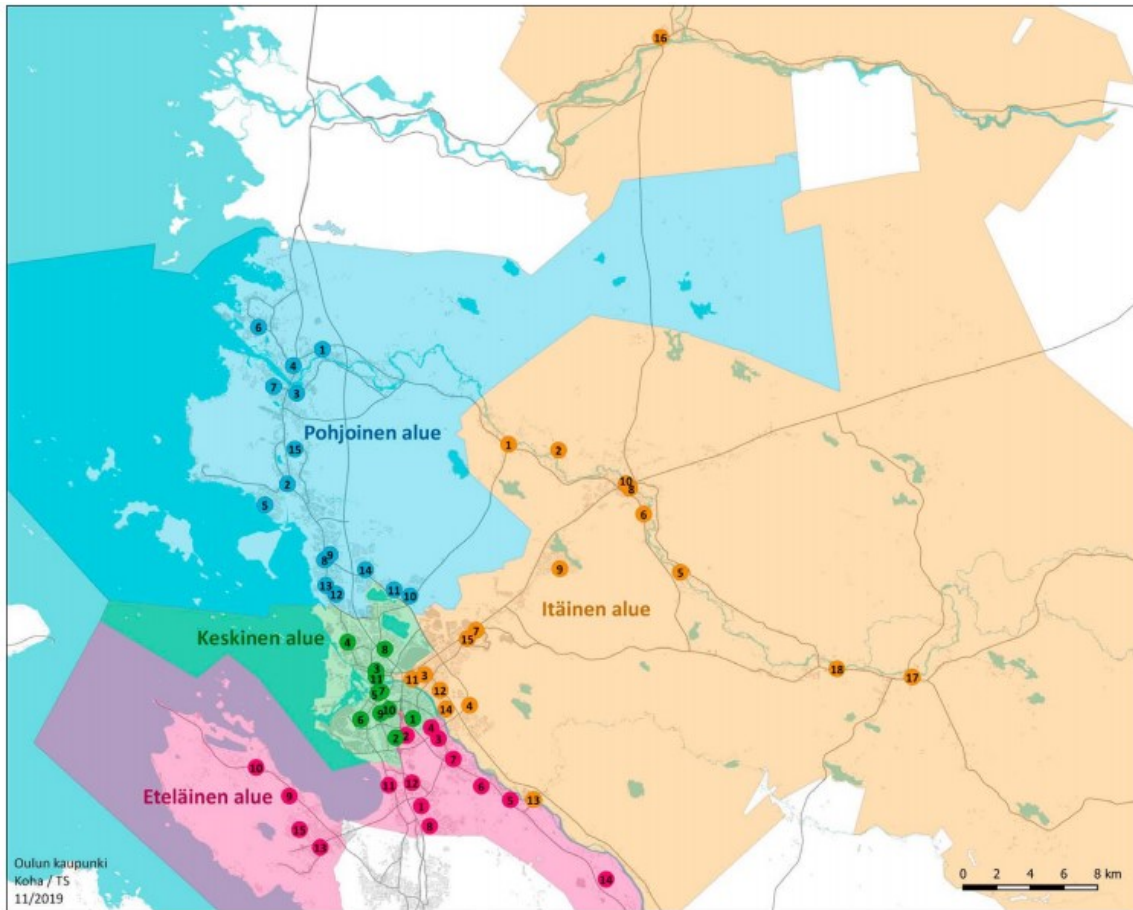
alueellisiin tekijöihin, eikä käytettävä YKR-aineisto sisällä valmiita jakoja tutkimuksessa käytettäviin alueisiin. YKR-aineistoa jouduttiin näin ollen käsittelemään sopivaksi alueellisia tarkasteluja varten rajaamalla ArcGIS-ohjelman avulla sopiviin eri kokoisiin alueisiin ja palasiin. Tutkielmaan käytettävissä ollut YKR-aineisto sisälsi lisäksi paljon erilaisia tutkimuksen toteuttamisen kannalta turhia tietoja. Esimerkkejä tällaisista tiedoista ovat muun muassa asuntokuntien asunnon omistamiseen liittyvät tiedot tai työvoiman tarkka alakohtainen erittely. Joissain tapauksissa esimerkiksi työvoiman alakohtaisesta jakautumisesta olisi hyötyä vastaavanlaisessa tutkimuksessa, mutta tässä tapauksessa tällaisen muuttujan käsittely olisi lisännyt tutkielman työläyttä merkittävästi tuottamatta kuitenkaan juuri lisäarvoa.

Tutkielman toteutuksessa on lisäksi otettava huomioon, että Oulun kaupungilla on käytössä neljä suurta oppilasaluetta, joiden sisällä oppilaat jakautuvat eri kouluihin. Suuria oppilasalueita, eli SIKU-alueita tarkempaa tietoa koulujen oppilasalueista ei tutkielman toteutuksessa ollut käytettävissä, joten tutkielmassa tyydytään sellaiseen oletukseen, että oppilaat käyvät pääosin kotiansa lähimpänä sijaitsevaa koulua. Oulun kaupungin oppilaaksioton periaatteiden mukaisesti näin ei kuitenkaan välttämättä aina ole. Kaupungin linjauksien mukaisesti ”lähikoulu” voi olla myös jokin muu kuin oppilaan kotia lähimpänä sijaitseva koulu (Oulun kaupungin sivistys- ja kulttuurilautakunta, 2019).

Oulun kaupungin oppilaaksioton periaatteista voidaan kuitenkin tehdä sellainen päätelmä, että ensisijaisesti lapset käyvät lähimpänä kotiaan sijaitsevaa koulua. Kaupungin linjaamien oppilaaksioton periaatteiden mukaisesti terveydentilaan liittyvän tai muun erityisen syyn puuttuessa seuraavaksi tarkastellaan sisarusperustetta. Mikäli oppilaalla ei ole sisarusia tai sisarus on kansainvälisessä koulussa, erikoisluokalla, päässyt hakemuksen perusteella johonkin muuhun kuin lähikouluun, tai on sijoitettu erityisen tuen päätöksellä toiseen kouluun, siirrytään tarkastelemaan koulumatkaperiaatetta. (Oulun kaupungin sivistys- ja kulttuurilautakunta, 2019.) Muuhun kuin lähikouluunsa hakevien ja päässeiden oppilaiden huomioiminen mutkistaisi aineiston käsittelyä niin merkittävästi, että tässä pro gradu -tutkielmassa ei ole mahdollisuutta ottaa tätä huomioon. Todennäköisesti määrät eivät kuitenkaan ole niin merkittäviä, että niillä olisi kovin suurta vaikutusta tutkimuksen tuloksiin. Näin ollen tutkimuksessa tehdään oletus, että tulosten kannalta merkittävin osa oppilaista tulee koulun lähialue-estimaatin sisäpuolelta.

3.2.1 Lähialue-estimaatit ja tietojen poimiminen

Koska Oulun kaupungissa on vain neljä oppilasaluetta ja Oulu on maantieteellisesti hyvin laaja kaupunki, ovat oppilasalueet hyvin suuria. Tämän vuoksi ei tutkimuksellisesti ole mieltä käyttää näitä hallinnollisia oppilasalueita kuvaamaan koulujen toimintaympäristöjä. Mikäli tutkimuksessa käytettäisiin koko oppilasalueen ominaisuuksia kuvaamaan koulujen lähialueita tai naapurustoja, katoaisivat koulujen toimintaympäristöjen eroavaisuudet koko oppilasalueen yleiseen keskiarvoon. Esimerkiksi Oulun kaupungin itäinen ja pohjoinen oppilasalue sisältävät hyvin erilaisia asuinalueita, sillä niiden alueilla sijaitsee sekä kaupunkiasutusta, että harvempaan asuttua seutua. Lisäksi erilaisilla asuinalueilla tai naapurustoilla kuten – yliopiston sijainnin vuoksi pitkälti opiskelijoiden asuttamalla Linnanmaalla – ja jollain toisella asuinalueella – kuten lapsiperheiden asuttamalla Ritaharjulla – voi muutenkin olla keskenään hyvinkin erilaiset ominaispiirteet. Kuvassa 1 on näkyvillä Oulun kaupungin oppilasalueet, jotka ovat myös Sivistys- ja kulttuurialueita (SIKU-alueita). Kuva auttaa hahmottamaan alueiden suuret koot, sekä mahdollisesti sen millaisia eroja esimerkiksi pelkässä kaupunkirakenteessa voi olla alueiden sisällä. Esimerkiksi Keskinen alue on suurilta osin hyvin tiiviisti asuttua kaupunkikeskustaa ja muita tiiviisti asuttuja alueita, kun taas Itäisellä alueella kaupunkimaista asutusta on lähinnä vain sen läntisimmässä reunassa, muun alueen ollessa pitkälti haja-asutusta ja joitakin pienempiä asutuskeskittymiä. Numerot kuvassa ovat kaupungin kouluja.

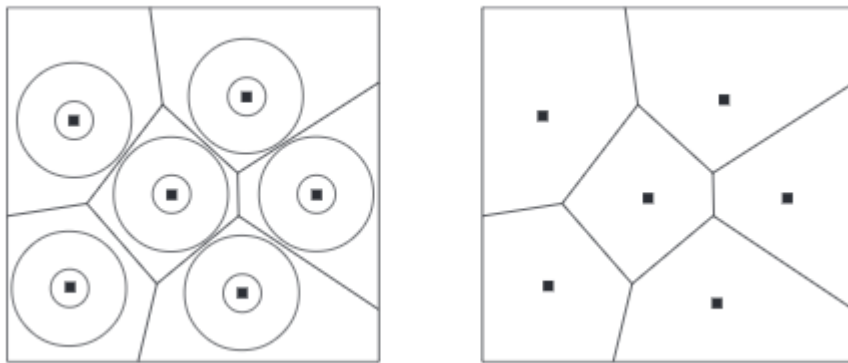


Kuva 1. Oulun kaupungin oppilasalueet (Kuvan lähde: Oulussa koulussa - perusopetuksen opas 2020–2021 (Oulun kaupunki, 2019.))

Aiemmissa tutkimuksissa on hyödynnetty vastaavassa tilanteessa niin kutsuttuja lähialue-estimaatteja. Esimerkiksi Berneliuksen (2015, 636–637) tutkimuksessa lähialue-estimaatteja hyödynnettiin Espoossa ja Vantaalla, joissa oppilasalueet ovat Oulun tapaan suuria, eivätkä hallinnolliset alueet näin ollen riittä kuvaamaan koulujen lähiympäristöjä riittävän tarkasti. Myös tämän tutkielman kohdalla tultiin samaan päätelmään siitä, että lähialue-estimaattien käyttäminen on järkevää suurten oppilasalueiden vuoksi. Lisäksi, jotta työllä olisi vertailukohtia suomalaisessakin tutkimuksessa, päädyttiin tutkimuksen toteuttamisessa käyttämään samankaltaista ratkaisua kuin aiemmissa tutkimuksissa. Aiemmissa tutkimuksissa käytettyjen metodien hyödyntäminen tässäkin tutkielmassa mahdollistaa vertailun näihin aiempiin tutkimuksiin, jolloin tutkielman tulokset eivät jää pelkästään yksittäisiksi havainnoiksi koulujen alueellisesta eriytymisestä Oulussa, vaan niitä voidaan verrata kansallisessa kehityksessä muihin kaupunkeihin, joissa samanlaista tutkimusta on tehty.

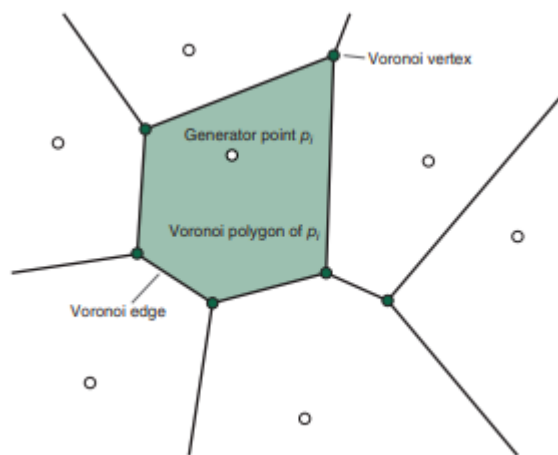
Berneliuksen tutkimuksessa (2015) lähialue-estimaatit tarkoittavat geometrisesti luotuja alueita, jotka ovat muodoltaan polygoneja eli monikulmioita. Näitä monikulmioita kutsutaan

yleisesti Thiessenin polygoneiksi, mutta usein myös Voronoin diagrammeiksi. (Longley, Goodchild, Maguire, & Rhind, 2005, 333). Thiessenin polygonit tunnetaan useimmiten amerikkalaisen meteorologin Alfred H. Thiessenin nimellä, sillä hän kehitti tälle jopa satoja vuosia tunnetulle konseptille käytännöllisen käyttötavan parantaakseen alueellisten sademäärien arviointia (Yamada, 2017, 1–2). Kyseisten polygonien muodostamista voidaan kuvata esimerkiksi siten, että toivotuista lähtöpisteistä ”puskuroidaan” (oma suomennos) yhtäaikaaisesti jatkuvasti joka suuntaan ulospäin laajeneva alue. Kun nämä eri pisteistä lähtevät ”puskurit” törmäävät, ne lopettavat laajenemisen tässä kohdassa ja jatkavat laajenemista muilla alueilla niin kauan, että kohtaavat uudelleen toisen alueen puskurin reunan. Tuloksena syntyy polygoneja eli monikulmioita kuvan 2 tapaan (Albrecht, 2007, 44.)



Kuva 2. Thiessenin polygonien muodostaminen. (Kuvan lähde: Albrecht 2007, 44).

Jos ”keskipisteet”, joita kutsutaan generaattoripisteiksi, ovat esimerkiksi joukko $P = \{p_1, \dots, p_n\}$, polygoni, ja useista polygoneista muodostuva Voronoin diagrammi muodostetaan osoittamalla jokaiselle sijaintipisteelle alueella lähin joukon P jäsen. Ne pisteet, jotka sijaitsevat yhtä kaukana kummastakin P :stä jaetaan/osoitetaan kummallekin. Nämä rajapisteet muodostavat Voronoin rajan (Voronoi edge), joka muodostaa polynomien rajan. Polynomin kulmissa rajat kohtaavat ja muodostavat niin sanotun Voronoin kärjen (Voronoi vertex). (Yamada, 2017, 2–3.)



Kuva 3. Voronoin diagrammin osat. Vihreän alueen pisteet sijaitsevat lähimpänä p_1 -pistettä. (Kuvan lähde: Yamada, 2017, 2.)

Albrechtin kuvaaman puskurointitavan lisäksi polygonien muodostamisen operaatio voidaan esittää myös siten, että kahden generaattoripisteen välille piirretään viiva. Tämän jälkeen tarkalleen tämän viivan puoliväliin piirretään 90 asteen kulmassa toinen viiva, joka siis ilmaisee pisteiden välisen keskikohdan. Kun sama operaatio toistetaan kaikkien eri pisteiden välillä (esimerkiksi generaattoripisteiden A , B ja C välillä: $A-B$, $A-C$, $B-C$), muodostuu viivoista lopulta monikulmioita, joissa niiden rajat muodostuvat aina kahden eri keskuksen puoliväliin. Mistä tahansa kohdasta polygonin sisällä on lyhempi matka kyseisen polygonin generaattoripisteeseen kuin mihinkään muuhun Voronoin diagrammissa sijaitsevaan generaattoripisteeseen (Yamada, 2017, 1).

Thiessenin polygoneja käytetään muun muassa arkeologian, kartografian, maantieteen, geologian, meteorologian, markkinoinnin, fysiikan, sekä kaupunki- ja aluesuunnittelun kentillä, sekä monella muullakin tutkimusalueella. Jokaisen polygonin kattama alue voidaan käsittää sen ”keskuksen” eli generaattoripisteen vaikutusalueeksi. Esimerkiksi pikaruokaravintoloiden ympärille luotuja polygoneja voidaan käyttää arvioimaan näiden ravintoloiden markkina-alueita, jos oletetaan että niiden tarjoamat palvelut ovat täsmälleen samanlaisia ja hintaisia, jolloin asiakkaiden on järkevintä käydä sijaintiaan lähimmässä ravintolassa. (Yamada, 2017, 1.) Thiessenin polygoneja käytetään yleisesti geoinformatiikassa nopeuttamaan joitakin operaatioita, kuten esimerkiksi lähimmän naapurin etsintää (Longley ym., 2005, 333).

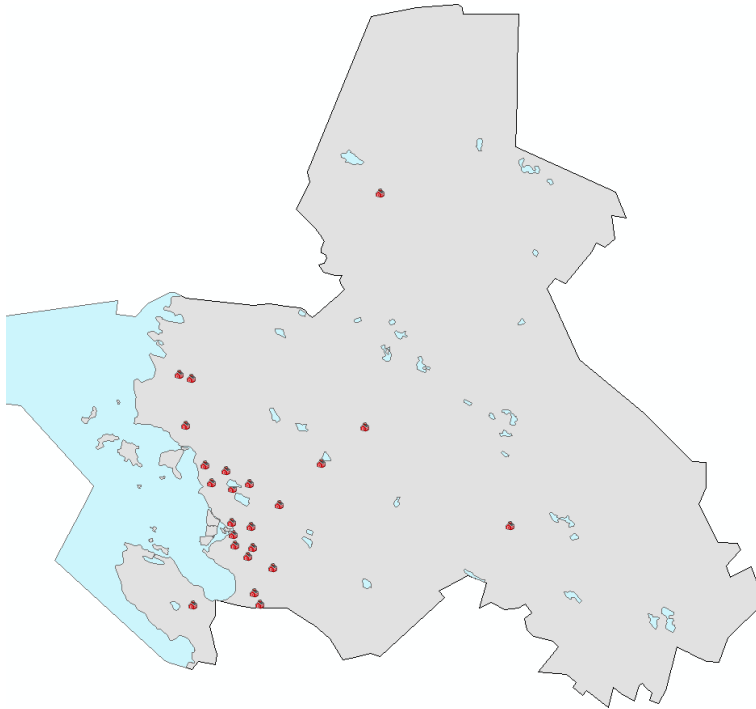
Oulun peruskoulujen ympärille luotavat Thiessenin polygonien voidaan katsoa näyttävän lähikouluperiaatteen mukaiset oppilasalueet jokaiselle koululle ainakin geometrisesti, sillä

jokaisen polygonin sisällä matka kyseisen polygonin sisällä olevaan kouluun on lyhempi kuin mihinkään muuhun kouluun. Eli jos Oulun peruskoulut ovat generaattorijoukko $P = \{p_1, \dots, p_n\}$, on jokaisesta mahdollisesta pisteestä polygonin sisällä lyhin matka kyseisen polygonin sisäiseen generaattoripisteeseen p_n , joka siis edustaa koulua. Samalla tavalla voidaan muodostaa koulujen ympärille niiden lähikouluperiaatteen mukaisten oppilasalueiden mukaiset naapurustot. Näistä naapurustoista voidaan sitten laskea tutkimuksessa tarvittavia tietoja jokaisen polygonin sisältä erikseen.

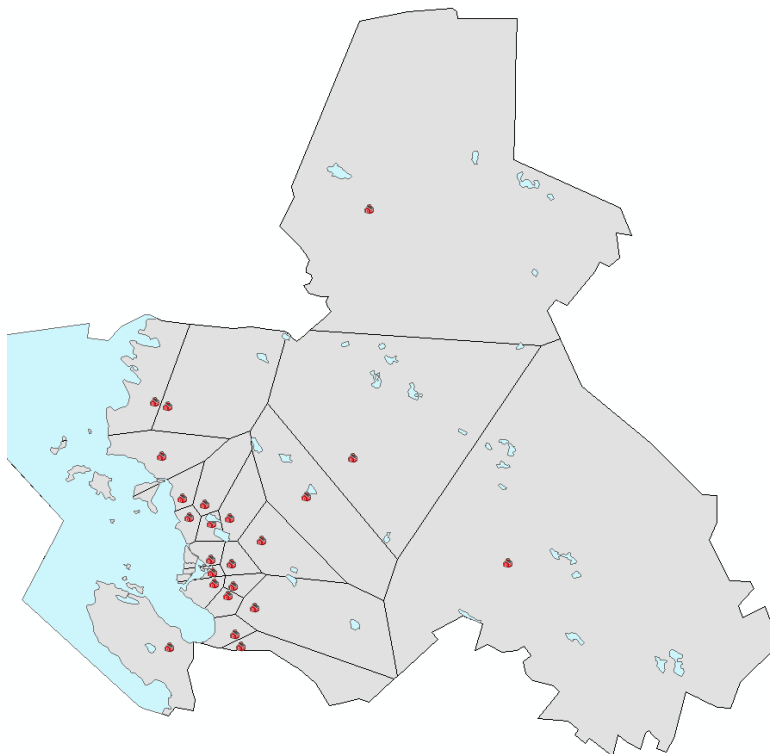
Thiessenin polygonit siis vastaavat hyvin tutkimuksen toteuttamisessa esiintyvään tarpeeseen muodostaa kouluille naapurustot tai lähialueet, joista saatavien tietojen perusteella on mahdollista verrata sosioekonomisesti erilaisilla alueilla sijaitsevien koulujen oppimistuloksia toisiinsa. Thiessenin polygonien muodostaminen on myös operaationa suhteellisen helppo ja nopea toteuttaa esimerkiksi ArcGIS -ohjelmassa, jota tässä tutkimuksessa käytettiin. Näin ollen lähialue-estimaattien muodostamiseen ei kulu tutkimuksessa suhteettoman paljon aikaa ja vaivaa, vaan aineistosta saadaan käyttökelpoisia havaintoja ulos melko lyhyessä ajassa.

Lähialue-estimaatit päädyttiin tutkimuksessa muodostamaan kahdella tavalla, jotta voitaisiin vertailla miten eri tavoin rajautuvat polygonit vaikuttavat tuloksiin. Ensimmäisessä tapauksessa muodostettiin lähialue-estimaatit siten, että Thiessenin polygonit luotiin kaikille mukana oleville kouluille samanaikaisesti, käyttäen käsittelyn laajuutta rajaavana tekijänä kaupungin hallintoalueen rajojen pohjoisinta, itäisintä, eteläisintä, sekä läntisintä pistettä. Näin polygonit muodostivat Voronoin diagrammin, joka peitti koko kaupungin hallintoalueen. Muodostunut Voronoin diagrammi leikattiin sitten ArcMapissa vastaamaan kaupungin rajoja. Apuna tässä toimenpiteessä käytettiin kaupungin hallintoalueen rajat sisältävää shapefile-tiedostoa, joka ladattiin Maanmittauslaitoksen avoimien aineistojen tietokannasta. Näin muodostettuun karttakuvaan lisättiin myös vesialueet kaupungin hallintoalueen sisäpuolelta.

Kuvissa merkitty myös tutkimuksessa käytettyjen yläkoulujen sijainnit koulua kuvaavalla symbolilla.

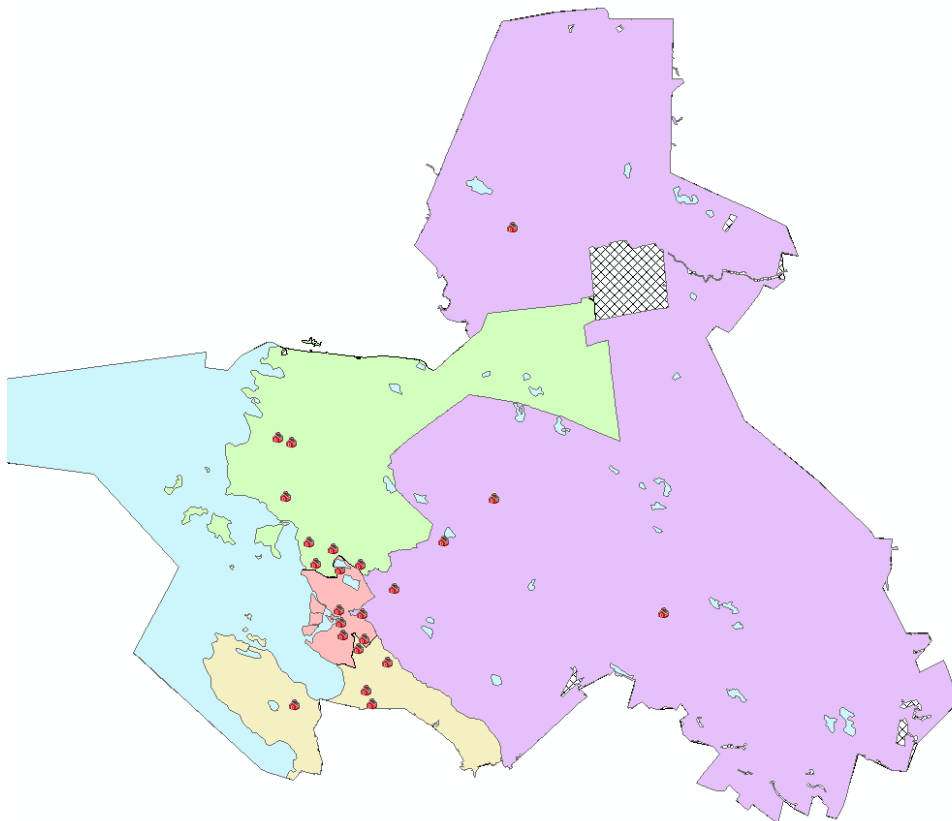


Kuva 4. Alkutilanne polygonien muodostamiseksi ilman SIKU-alueita.

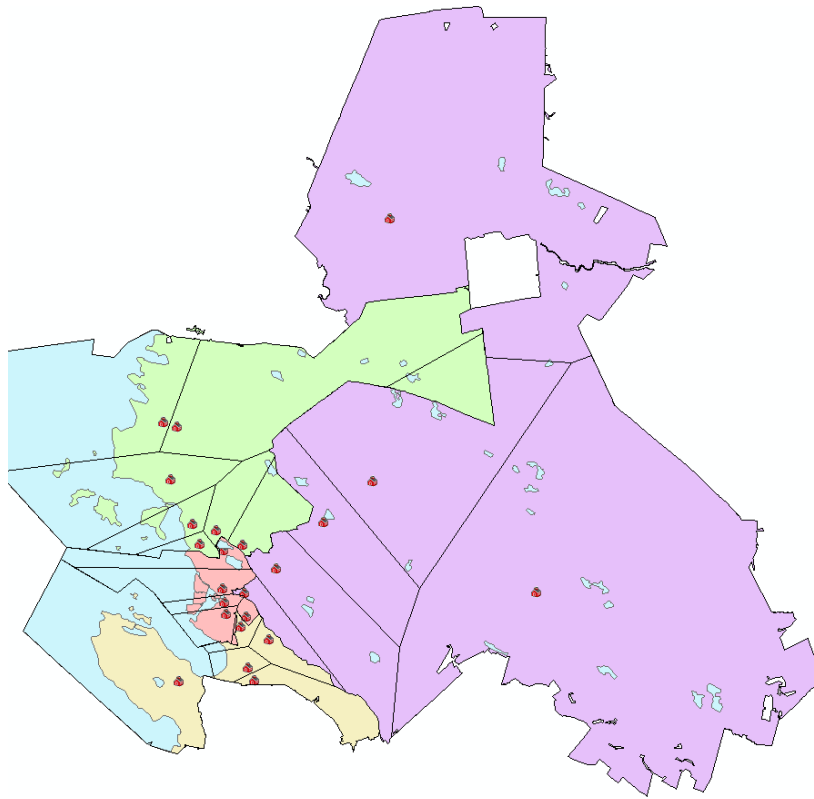


Kuva 5. Lähialue-estimaatit muodostettuina siten, että ne rajautuvat joko toiseen polygoniin tai Oulun kaupungin rajaan.

Toisessa tapauksessa lähialue-estimaatit muodostettiin ja käsiteltiin muutoin samoin kuin ensimmäisessä, mutta polygonien muodostumista rajasivat kaupungin rajojen sijaan yksittäisen SIKU-alueen rajat. Polygonit siis muodostettiin käsittelemällä vain tietyllä SIKU-alueella sijaitsevia kouluja, antaen polygonien muodostua kyseisen SIKU-alueen rajojen puitteissa, minkä jälkeen alueiden rajojen ulkopuoliset osat rajattiin pois käyttäen Oulun kaupungin avoimen datan osastolta saatuja SIKU-alueiden rajoja. Näin muodostettu kartta nähtävissä alapuolella. Eri värein kuvatut alueet ovat eri SIKU-alueita.



Kuva 6. Alkutilanne SIKU-alueiden rajoihin rajautuvien lähialue-estimaattien muodostamiseksi.



Kuva 7. Lähialue-estimaatit muodostettuina siten, että ne rajautuvat SIKU-alueiden rajoihin ja kaupungin rajaan.

Kun lähialue-estimaatit oli kahdella kuvatulla tavalla muodostettu, yhdistettiin luotuihin karttoihin ArcMapissa myös paikkatietoruutuihin sidottu YKR-aineisto. YKR-aineiston sisältävien paikkatietoruutujen lisäämisen jälkeen poimittiin ArcMapin valintatyökalua käyttäen erikseen jokaisen polygonin sisällä sijaitsevat YKR-aineiston tiedot sisältävät ruudut, sekä kirjattiin ylös valittujen ruutujen määrä ja tutkimuksessa käytettävien tietojen summa kyseisistä ruuduista. Valinnassa käytettiin asetusta, joka poimii jokaisen 250x250 metriä kokoisen paikkatietoruudun sen alueen sisältä, johon ruudun painopiste asettuu. Näin välttyttiin poimimasta osittain polygonien rajojen päällä sijaitsevia paikkatietoruutuja kahteen tai useampaan lähialue-estimaattiin mukaan. Lisäksi rajattiin pois ne tilastoruudut, joista poimittavan rahatulon summa oli tietoturvasyistä salattu merkitsemällä tietoihin lukema -1, sillä mukaan otettaessa näiden ruutujen väestö- ja asukaskuntamäärät olisivat vääristäneet rahatulolaskelmia. Näiden ruutujen tulo- tai asukastiedot eivät siis ole mukana tutkielmassa. Kyseisten -1-ruutujen kohdalla tehdystä ratkaisusta seuraa joitakin ongelmia, joita käsittelem myöhemmin useampaankin otteeseen.

3.2.2 Muuttujien valinta ja muodostaminen

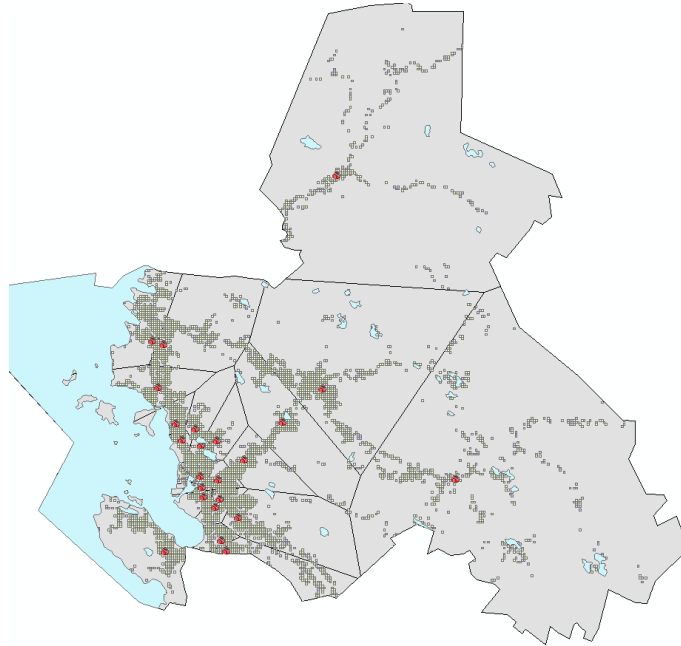
Muuttujien muodostamisessa lähdettiin liikkeelle siitä, että taustateorian pohjalta pyrittiin löytämään merkittäviä ja vastaavassa tutkimuksessa aiemmin käytettyjä muuttujia, jotta tuloksia voitaisiin vertailla. Taustateoriaan nojautuen haluttiin käyttää muuttujia, jotka kuvaisivat lähialueiden keskimääräistä tulo- ja aikuisväestön koulutustasoa.

Tiedot poimittiin ruututietokannasta ArcMap -ohjelman avulla. Tilastoruuduista poimitut tiedot merkittiin Exceliin kyseisen lähialue-estimaatin oman yläkoulun kohdalle, jotta tietoja olisi helppo käsitellä SPSS-analyysiohjelmassa, jossa analyysit suoritettiin. Ylös merkittiin myös se, kuinka monta tilastoruutua miltäkin alueelta pystyttiin poimimaan tietoturva/yksilönsuojan takia salattujen ruutujen poistamisen jälkeen. Lukema merkittiin ylös sekä muodossa **poimitut ruudut/ruutujen kokonaismäärä alueella**, että prosenttiosuutena pyöristettynä lähimpään kymmenykseen. Esimerkiksi näin: $78/98 = 79,6 \%$. Näin pystyttiin myöhemmässä vaiheessa arvioimaan kuinka hyvin aineisto kuvaa alueen todellista tilannetta.

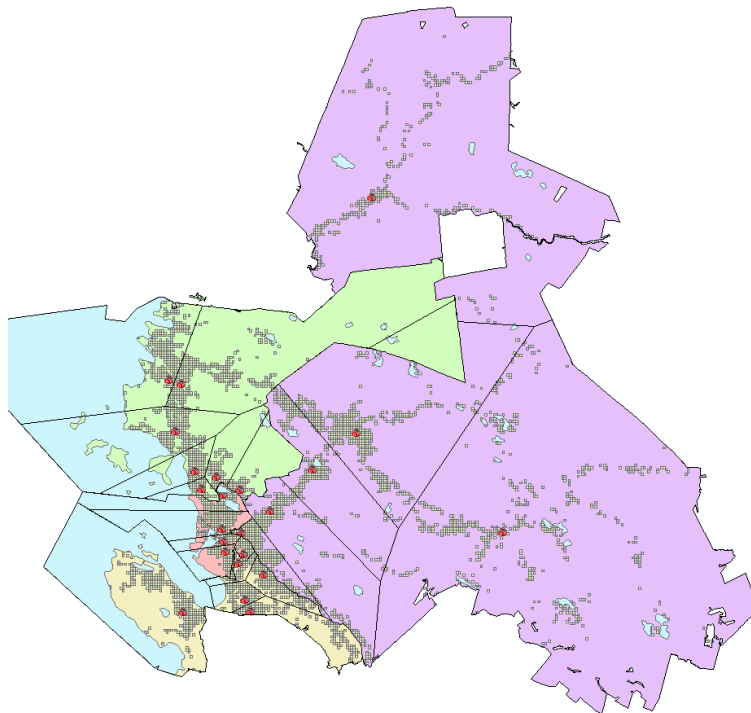
Aikuisväestön tulotason käyttäminen perustuu siihen, että hyvin toimeentulevien perheiden tiedetään olevan paremmassa asemassa esimerkiksi asuntomarkkinoilla, mikä voisi ennustaa näiden perheiden ”pakkautumista” tietyille alueille. Lisäksi vanhempien tulojen tiedetään vaikuttavan muun muassa painotetuille luokille hakeutumiseen, ja kun tiedetään painotetussa opetuksessa saavutettavan parempia tuloksia (esim. Kalalahti ym., 2015a, 23), voidaan ajatella, että koulujen oppimistulokset olisivat esimerkiksi tätä kautta yhteydessä lähialueen tuloihin. Voi nimittäin olla, että hyvätuloiset ja koulutetut, aktiivisesti kouluvalintaa tekevät vanhemmat käyttävät taloudellisia resurssejaan hyväksi ja muuttavat hyväksi koetun tai painotettua opetusta tarjoavan koulun alueelle - vaikka koulun koettu hyvä taso ei nousekaan vanhempien listoilla aivan muuttokriteerien kärkeen - sillä koulumatkojen pituutta ja alueen turvallisuutta pidetään merkittävinä tekijöinä asuntoa hankittaessa (Dhalmann ym., 2013). Voidaan olettaa, että hyvämaineisella ja turvallisella alueella sijaitsee myös vastaavia ominaisuuksia omaava koulu.

Tulomuuttujaa muodostettaessa tiedoista laskettiin **rahatulo per henkilö** (merkitty *rahatulo/henk*) jakamalla halutun lähialue-estimaatin tilastoruutujen rahatulon yhteissumma asukasmäärällä. Vastaavaa aluekohtaista tietoa keskimääräisestä tulotasosta on hyödynnetty aiemminkin (vrt. esim. Bernelius, 2013a, 192). Lisäksi haluttiin tutkia vaikuttaako rahatulon ja oppimistulosten väliseen yhteyteen, jos tulotasoa mitattaisiinkin perhekohtaisesti. Tämän vuoksi laskettiin **rahatulo per asutokunta** (*rahatulo/ak*) jakamalla lähialue-estimaatin

rahatulon kokonaissumma asutokuntien määrällä. Näille erilaisille rahatuloa alueella kuvaaville muuttujille laskettiin keskiarvo, -hajonta, sekä mediaani.



Kuva 8. Koulujen lähialue-estimaatit ilman oppilasalueiden rajoja + YKR-aineisto.



Kuva 9. Koulujen lähialue-estimaatit oppilasalueiden rajojen kanssa + YKR-aineisto.

Muuttujaksi haluttiin alueen tulotasoa kuvaavien muuttujien lisäksi myös alueen aikuisväestön koulutustasoa kuvaava muuttuja, sillä esimerkiksi Berneliuksen (2005, 47–48; 2013a, 178–179) tutkimustulosten mukaan koulujen lähialueiden aikuisväestön koulutustaso korreloi voimakkaasti sekä ala-, että yläkoulujen oppimistulosten kanssa. Lisäksi lukuisat tutkimukset ovat osoittaneet, että sosioekonomisista taustatekijöistä nimenomaan vanhempien koulutustaso ennustaa parhaiten lapsen koulumenestystä Suomessa (esim. Bernelius, 2005; 2013a; Hautamäki, Kupiainen, Marjanen, Vainikainen, & Hotulainen, 2013; Kalalahti, 2012; Kuusela, 2010; Lehti & Laaninen, 2020; Salmela-Aro & Chmielewski, 2019; Vainikainen ym., 2016; Välijärvi ym., 2015).

Tilastoaineistosta poimittiin siis tulotietojen lisäksi myös tiedot lähialue-estimaattien työssäkäyvän työväestön koulutusasteista. Työväestön koulutusasteeseen päädyttiin siitä syystä, että kyseinen työvoimatietojen yhteydestä löytyvä tieto oli ainoa käyttökelpoinen alueen asukkaiden koulutustasoa kuvaava muuttuja aineistossa. Koulutustasot oli ilmoitettu aineistossa siten, että perusasteen, keski- ja korkea-asteen koulutuksen saaneet työlliset oli kaikki eroteltu omiin luokkiinsa. Nämä koulutustasokohtaiset määrät poimittiin erikseen jokaisen lähialue-estimaatin kohdalla. Lisäksi poimittiin talteen jokaisen lähialue-estimaatin työllisen työvoiman määrä. Aineistossa oli ilmoitettu myös työvoiman määrä, mutta työvoiman määrä sisälsi paljon sellaisiakin henkilöitä, joiden koulutustasosta ei pystynyt päättämään mitään, joten siitä päätettiin luopua ja käyttää sen sijaan työllisten määrää. Tiedoista laskettiin kunkin koulutusasteen käyneen työvoiman osuus koko työllisten määrästä alueella, kuten esimerkiksi: **korkea-asteen koulutuksen saaneet työlliset/työllisten kokonaismäärä**. Saatua luku muutettiin prosenteiksi ja näin saatiin selville kunkin koulutusasteen saaneiden työllisten suhteellinen osuus kunkin alueen työväestöstä. Vastaavaa menetelmää on jälleen käyttänyt esimerkiksi Bernelius (2005, 47–48; 2013a, 178–179), joten vertailukohta on löydettävissä, vaikkakin Bernelius on käyttänyt esimerkiksi korkeakoulututkinnon suorittaneiden 25–64-vuotiaiden osuuksia koko alueen asukasmäärästä.

Myös koulutusmuuttujia muodostettaessa jouduttiin osa havaintoruuduista jättämään pois tietoja kerätessä, sillä työvoiman määrän ollessa ruudussa alle 10 olivat muun muassa koulutusastetiedot tulotietojen tapaan salatut luvulla -1. Salattujen arvojen mukaan ottaminen olisi tehnyt koulutustasojen suhteen työllisiin muodostamisen ja vertailun hankalaksi, joten tiedot poimittiin vain niistä ruuduista joissa **työvoima yhteensä > -1**. Näin ollen alueiden koko *työvoima* ei ole mukana laskelmissa, eivätkä myöskään kaikki *työlliset*, mutta sen sijaan mukana ovat kaikki työlliset joiden koulutustaso tiedetään eli *työssäkäyvä työvoima*. Näin ollen

muuttujat *tyol_per/tyol* (vain peruskoulun käyneet), *tyol_kes/tyol* (keskiasteen koulutuksen saaneet) ja *tyol_kor/tyol* (korkeakoulun käyneet) kuvaavat työssäkäyvän työväestön koulutustasojen suhteellisia osuuksia alueiden työssäkäyvistä työvoimasta prosenttiosuuksina ilmaistuna. Kun koulutusta kuvaavat muuttujat lasketaan yhteen, tulokseksi saadaan 100 % alueen työssäkäyvän työvoimasta.

Tiivistetysti YKR-aineiston käsittelyssä tehtiin ArcGIS-ohjelmassa seuraavat toimenpiteet, jotta aineisto saatiin tutkielman toteuttamisen kannalta sellaiseen muotoon, joka mahdollisti analyysin:

1. Paikkatietomuotoiset koulujen sijainnit yhdistettiin pohjakarttaan.
2. YKR-aineisto yhdistettiin pohjakarttaan.
3. Sivistys- ja kulttuurialueiden rajat yhdistettiin pohjakarttaan.
4. Koulujen ympärille muodostettiin Thiessenin polygonit/Voronoin diagrammi perustuen koulujen sijainteihin koordinaatistossa, sekä Oulun kaupungin hallintoalueen ja SIKU-alueiden rajoihin.
5. YKR-aineisto rajattiin Thiessenin polygonien avulla osiin niin, että jokaiselle YKR-aineiston ruudulle osoitettiin lähialue-estimaatti niiden sijaintipolygonin perusteella. Rajatapaukset määriteltiin siihen polygoniin, jonka rajojen sisälle niiden painopiste osui.
6. Lähialue-estimaateille osoitettujen YKR-ruutujen perusteella poimittiin kyseisen polygonin väestön tulo- ja koulutustiedoista analyysin kannalta tarpeelliset tiedot.
7. Aineisto oli käsitelty ja voitiin siirtyä varsinaiseen analyysivaiheeseen.

3.3 Analyysimenetelmät

Analyysissä tutkitaan, löytyykö muodostettujen oppilasalueiden tulo- ja työssäkäyvien koulutustasosta tilastollisesti merkittäviä yhteyksiä Oulun kaupungin toimittamiin yläkoulujen päättökeskiarvoihin. Yhteyksiä tutkitaan tilastollisin menetelmin korrelaatio- ja regressioanalyysien keinoin. Seuraavissa kahdessa alaluvussa esitellään käytetyt menetelmät tarkemmin.

3.3.1 Korrelaatioanalyysi

Korrelaatio ilmaisee kahden muuttujan välistä yhteyttä ja on tämän yhteyden tärkein indikaattori. Muuttujat voivat olla yhteydessä toisiinsa positiivisesti tai negatiivisesti. Positiivinen yhteys eli korrelaatio nähdään silloin, kun toisen muuttujan arvon kasvaessa myös toisen muuttujan arvo kasvaa. Korrelaatio eli muuttujien välinen yhteys voi olla myös negatiivinen, jolloin toisen muuttujan arvon kasvaessa toisen muuttujan arvo puolestaan laskee. Voi olla myös niin, että muuttujien välillä ei ole yhteyttä, jolloin muuttujat eivät korreloi keskenään. (Akoglu, 2018, 92; Metsämuuronen, 2011, 364; Schober, Boer, & Schwarte, 2018, 1763.)

Korrelaatiosta puhuttaessa on tärkeää muistaa, että vaikka muuttujien välillä näyttäisikin olevan yhteys, ei tämä kuitenkaan aina tarkoita *kausaliteettia*, eli syy-seuraus -suhdetta muuttujien välillä (Jokivuori & Hietala, 2014). Esimerkiksi kuuluisassa, usein käytetyssä esimerkissä jäätelön syönnin ja hukkumiskuolemien määrän välillä on korrelaatiota: kun jäätelöä syödään enemmän, on myös hukkumiskuolemia enemmän. Tästä voisi päätellä, että jäätelön syöminen aiheuttaa hukkumiskuolemia, tai toisinpäin. Todellisuudessa kyse lienee kuitenkin siitä, että jäätelöä syödään kesällä, jolloin myös uidaan enemmän, eli hukkumiskuolemat ja jäätelönsyönnin määrä riippuvatkin kolmannesta muuttujasta; siitä onko kesä vai ei. Tässä kuuluisassa esimerkissä jäätelönsyönnin ja hukkumiskuolemien välillä vallitsee korrelaatio, mutta ei kuitenkaan syy-seuraus -suhde eli kausaliteetti. On siis huomattava, että pelkkä korrelaation olemassaolo ei varsinaisesti kerro meille mitään, vaan tärkeää olisi tietää millä tavalla havaittu korrelaatio syntyy. Ei riitä, että todetaan asioilla olevan yhteys, vaan usein olisi tärkeää tietää kumpi seuraa kummasta (Metsämuuronen, 2011, 410), vai vaikuttaako taustalla kenties vielä jokin muu tekijä.

Korrelaatioanalyysia tehtäessä on tärkeää tietää, millaisilla mitta-asteikoilla käytettäviä muuttujia ilmaistaan. Syynä tähän on, että erilaiset korrelaatiota ilmaisevat tunnusluvut ja korrelaatiokertoimet edellyttävät tiettyntyyppisiä asteikkoja käytetyiltä muuttujilta (Metsämuuronen, 2011, 368). Seuraavana esitellään tutkimuksessa käytettyjen muuttujien mitta-asteikkoja.

Käytettävistä muuttujista keskitulo per asukas ja keskitulo per asuntokunta ovat *suhdelukuasteikollisia*. Suhdelukuasteikollisille muuttujille on ominaista, että tiedetään etäisyyden muuttujan arvojen X ja Y välillä olevan tietyn mittainen jollain yksiköllä mitattuna. Tutkimuksen kontekstissa tämä tarkoittaa esimerkiksi, että voidaan sanoa kuvitteellisten

keskitulojen 10 000 € ja 15 000 € välillä olevan 5000 euron ero tai etäisyys. Mittayksikkönä tässä tapauksessa on euro ja aina lisättäessä summaan yksi euro liikutaan sama matka rahan tai tulojen määrässä. Rahamäärät ja niiden väliset etäisyydet ovat siis tarkasti mitattavissa. Kun etäisyys voidaan tarkasti mitata jonkin yksikön avulla, kyseessä on vähintään välimatka-asteikollinen muuttuja. Suhdeasteikollinen muuttuja puolestaan on kyseessä, kun tiedetään että jossain pisteessä mitattava ominaisuus loppuu kokonaan. (Metsämuuronen, 2011, 70–72, 368; Valli, 2015.) Tutkimuksen tapauksessa tulojen kohdalla näin voidaan todeta; raha/tulot loppuvat pisteessä 0 euroa.

Kolmantena selittävänä muuttujana tutkimuksessa on koulutustaso. Koulutustaso itsessään on *järjestysasteikollinen* muuttuja sillä perusasteen, keskiasteen ja korkea-asteen koulutukset voidaan asettaa ”paremmuuden” tai ”suuruuden” perusteella järjestykseen. Ei voida kuitenkaan mitata esimerkiksi, että korkea-asteen koulutus olisi vaikkapa 3 kertaa parempi kuin keskiasteen. Muuttujista voidaan myös tunnistaa jakauman keskipiste, mutta aritmeettista keskiarvoa näille ei kuitenkaan voida laskea. (Metsämuuronen, 2011, 69, 366; Tilastokeskus, n.d.; Valli, 2015.) Tutkimuksessa koulutustasoja ei kuitenkaan käsitellä järjestyks- eli ordinaaliasteikollisina muuttujina, vaan tutkimuksessa hyödynnetään tietoja työväestön koulutusasteiden suhteellisesta määrästä lähialue-estimaatin sisällä. Koulutustasoja siis tarkastellaan tutkimuksessa prosenttiosuuksina alueen koko työväestöstä. Tarkasteltaessa koulutusasteiden suhteellisia osuuksia prosentteina koulutusasteiden tarkastelu tapahtuukin itseasiassa siten, että koulutusastetta voidaan käsitellä suhdeasteikollisena muuttujana. Voidaan nimittäin todeta, että jos alueen työssäkäyvistä työväestöstä 80 % on korkeakoulutettuja, se on 10 prosenttiyksikköä enemmän kuin 70 %. Etäisyys suhteellisten osuuksien välillä on tarkasti mitattavissa, minkä vuoksi mitta-asteikko ilmaisee välimatkaa. Lisäksi voidaan todeta, että prosenttiosuuksilla on myös tarkka nollakohta, jossa kyseinen ominaisuus lakkaa olemasta, jolloin kyseessä on suhdeasteikollinen muuttuja.

Tutkimuksessa käytettävistä muuttujista päättökeskiarvo on hieman haasteellisempi tapaus. Päättökeskiarvo itsessään on numeraalinen, minkä lisäksi se on myös jatkuva. Numero voidaan kuitenkin muuttaa myös arvosanaksi, jolloin se voidaan tulkita myös *välimatka-asteikolliseksi muuttujaksi*. Järjestyks- eli ordinaaliasteikollinen muuttuja pystyy ilmaisemaan, onko jotain ominaisuutta enemmän tai vähemmän, mutta se ei kuitenkaan pysty kertomaan sitä, kuinka paljon enemmän tai vähemmän kyseistä ominaisuutta on (Metsämuuronen, 2011, 366; Valli, 2015). Ordinaaliasteikollisia muuttujia voisivat olla esimerkiksi armeijan arvojärjestys tai koulutustasot. Arvosana voi olla esimerkiksi ”hyvä” tai ”erinomainen”, mitkä pystytään

järjestelemään mielekkääseen järjestykseen. Kouluarvosana voidaan käsittää kuitenkin olla myös numeerisena muuttujana, josta voidaan laskea esimerkiksi keskiarvoja. Esimerkiksi arvosana ”hyvä” merkitsee numeroa 8 todistuksessa. Järjestysasteikolliset muuttujat eivät kuitenkaan usein ole numeerisia ja kouluarvosanaa voidaankin pitää jonkinlaisena erityistapauksena niiden joukossa, sillä se voidaan ilmaista myös numeerisesti (KvantiMOTV, 2007). Kouluarvosana on sikäli erikoinen numeerinen muuttuja, että siihen ei liity mitään tarkkaa mitta-asteikkoa, kuten esimerkiksi vuosi tai tunti ajan ollessa mittauksen kohteena. Ei voida esimerkiksi väittää, että kaikissa luokissa, puhumattakaan kaikista kouluista, saa saman arvosanan samalla todellisella osaamisella. Ero voi olla hyvinkin suuri eri ryhmien välillä, olivat ryhmät sitten luokkia tai kouluja. Esimerkiksi numeroiden 4–5 ja 7–8 välinen yhden numeron askel ei todennäköisesti todellisuudessa ole yhtä suuri osaamisen määrässä.

Päättökeskiarvon järjestysasteikollisuutta voidaan puolustaa esimerkiksi toteamalla, että luokissa yhden opettajan pitämistä kokeista voidaan väittää, että ne yleensä noudattavat ainakin jollain tasolla jonkinlaista välimatka-asteikollisuutta. Kuitenkin jo siirryttäessä tasolle, jossa arvioitsijoita (arvioita antavia opettajia) on useampia, putoaa välimatka-asteikollisuus hyvin tulkinnanvaraiseksi, mikä puhuu välimatka-asteikollisuutta vastaan. Näin ollen numerot eivät välttämättä todellisuudessa ilmaise osaamista välimatka-asteikolle sopivalla eksaktilla tavalla, vaan päättönumero sopisi pikemminkin olemaan järjestysasteikollinen muuttuja. Koska käytössä nimenomaan päättökeskiarvo numerona, päädyttiin tutkimuksen toteutuksessa kuitenkin käsittelemään päättökeskiarvoa jatkuvana ja välimatka-asteikollisena muuttujana ottamatta kantaa sen todelliseen olemukseen oppimisen mittarina. Tämä mahdollistaa regressioanalyysin mielekkään käytön, sillä eräänä lineaarisen regressioanalyysin ehtona on, että käytettävä riippuva muuttuja on ja vähintään välimatka-asteikollinen kuten myöhemmin esitetään.

Koska kaikki käytetyt muuttujat ovat vähintään välimatka- eli *intervalliasteikollisia*, tutkimuksessa päätettiin lähtökohtaisesti muuttujien välisten yhteyksien kuvaamiseksi Pearsonin tulomomenttikorrelaatiokerrointa r , jota yleisesti nimitetään vain Pearsonin korrelaatioksi tai korrelaatioksi (Metsämuuronen, 2011, 369, 577; Valli, 2015). Jotta Pearsonin tulomomenttikorrelaatiota voitaisiin käyttää muuttujien välisten yhteyksien analysointiin, on muutamien ennakkoehtojen täytyttävä: Ensinnäkin molempien muuttujien on oltava vähintään välimatka-asteikollisia ja jatkuvia (vaikkakin Pearsonin korrelaatiokerroin on todellisuudessa tulkittavissa jo hyvällä järjestysasteikolla mitattujen muuttujien väliltä (Metsämuuronen, 2011, 577).), niiden täytyy olla normaalisti jakautuneita, muuttujien on oltava lineaarisessa

yhteydessä, muuttujien arvot eivät saa olla toisistaan riippuvaisia, sekä lisäksi aineistossa ei saa olla oleellisia outlier-tapauksia, eli poikkeavia havaintoja. (Schober ym., 2018, 1764–1765.) Jos näitä ennakkoehtoja rikotaan, on harkittava jonkin toisen korrelaation tutkimisen menetelmän käyttöä, kuten Spearmanin järjestyskorrelaatiokertoimen. Myöhemmin analyysiluvussa käykin ilmi, että tutkielman aineistossa parempi yhteyden mittari onkin todennäköisesti Spearmanin järjestyskorrelaatiokerroin.

Kumpikin edellä esitetyistä korrelaation mittareista voi saada arvoja välillä -1 , $+1$ ja mitä lähempänä nollaa arvo on, sitä heikompi on yhteys muuttujien välillä, numeron 1 puolestaan merkitessä täydellistä riippuvuutta. Arvojen vaihtelua välillä miinus yhdestä yhteen voidaan kuvata myös sanallisesti. Jos korrelaatiokertoimen arvo vaihtelee välillä $0,60$ – $0,70$ sitä voidaan kuvailla ”erittäin korkeaksi”, $0,60$ – $0,80$ välillä ”korkeaksi”, $0,40$ – $0,60$ välillä melko korkeaksi tai ”kohtuulliseksi”. Rajat eri sanallisten kuvailujen välillä kuitenkin vaihtelevat riippuen esimerkiksi tieteenalasta (kts. esim. Akoglu, 2018, 92). Ihmistieteissä harvoin päästään yli $0,80$ korrelaatioihin. Korrelaatiokertoimen etumerkki ilmaisee yhteyden suunnan, joka siis voi olla joko positiivinen, jolloin toisen ominaisuuden lisäys ennustaa toisenkin lisäämistä, tai negatiivinen, jolloin toisen ominaisuuden lisääntyminen vähentää toisen määrää tai päinvastoin. (Akoglu, 2018, 92; Metsämuuronen, 2011, 370–371; Schober ym., 2018, 1763; Valli, 2015.)

3.3.2 Regressioanalyysi

Regressioanalyysi on tilastollinen analyysimenetelmä, jonka avulla haetaan parhaalla mahdollisella tavalla selitettävän eli riippuvan muuttujan ennustavaa riippumattomien muuttujien (selittävät muuttujat) yhdistelmää. Regressioanalyysin lähtökohtana ovat välimatka- ja suhdeasteikon tasoiset muuttujat, mutta sen avulla voidaan tutkia myös järjestys- ja normaaliasteikollisia muuttujia, kunhan ne käsitellään dummy-muuttujiksi. (Heikkilä, 2014, 222.) Regressioanalyysi lukuisine muunnelmineen on tieteellisessä tutkimuksessa kenties yksi käytetyimpiä ja keskeisimpiä analyysimenetelmiä haluttaessa mallintaa todellisuutta analyysin keinoin. Perinteisessä muodossaan regressioanalyysi on yksinkertaistettuna menetelmä, jossa usean muuttujan avulla selitetään yhtä muuttujaa. Regressiomallin erityisenä etuna korrelaatiokertoimiin verrattuna voidaan pitää, että sen avulla voidaan tutkia kahden tai useamman selittävän muuttujan vaikutuksia selitettävään muuttujaan yhtä aikaa. Regressioanalyysijä voidaankin pitää yhteiskunta- ja ihmistieteissä monimuuttujamenetelmien

kulmakivenä, sillä näissä tieteenaloissa selittäviä tekijöitä on usein yhtä aikaa useampia kuin yksi. (Jokivuori & Hietala, 2014; Metsämuuronen, 2011, 709–711.)

Menetelmänä regressioanalyysi soveltuu joko tutkittavan ilmiön kannalta oleellisten muuttujien etsimiseen tai teorian kannalta oleellisten muuttujien vaikutuksien tarkasteluun. Tämä tarkoittaa, että regressioanalyysia voidaan käyttää joko tutkivasti (*eksploratiivisesti*) tai *konfirmatorisesti* eli selittävästi. Eksploratiivisessa tavassa selittävien muuttujien joukkoon valitaan aluksi useita tekijöitä, joista pyritään analyysin avulla erottelemaan tutkittavan ilmiön kannalta oleellisimmat. Toisessa, eli konfirmatiivisessa tavassa puolestaan yritetään jo alusta alkaen valita mukaan vain teorian kannalta oleellisimmat muuttujat, jonka jälkeen regressioanalyysin avulla pyritään vahvistamaan teorian kannalta keskeisten muuttujien selitysvaima. Regressioanalyysin avulla voidaan siis etsiä laajasta muuttujajoukosta juuri niitä tekijöitä, jotka selittävät parhaiten riippuvaa muuttujaa. Konfirmatiivisessa muodossaan – joka on Metsämuuronen mukaan regressioanalyysin sivistyneempi muoto – regressioanalyysin avulla ”*saatetaan tutkia jo aiemmin tärkeiksi tiedettyjen muuttujien osuutta selittävinä tekijöinä.*” (Jokivuori & Hietala, 2014; Metsämuuronen, 2011, 711–712.)

Lisäksi regressioanalyysia voidaan käyttää ilmiön mallintamiseksi ja mallinnuksen kautta myös uusien – aineistoon kuulumattomien – havaintojen ennustamiseen. (Metsämuuronen, 2011, 711.) Tämän tutkimuksen kontekstissa voitaisiin ajatella, että jos esimerkiksi onnistutaan mallintamaan onnistuneesti alueen aikuisväestön tulojen vaikutusta alueella sijaitsevan yläkoulun oppimistuloksiin, voitaisiin luodun mallin perusteella ennustaa tietylle alueelle perustettavan yläkoulun oppilaiden oppimistuloksia. Tässä tutkimuksessa ei kuitenkaan yritetä ennustaa uusia havaintoja, vaan mielenkiinto kohdistuu ainoastaan olemassa oleviin havaintoihin.

Mikäli käytettävästä tilastoaineistosta löytyy muuttujien väliltä säännönmukaisuuksia, voidaan siirtyä tutkimaan sitä, millaisen mallin avulla niitä voitaisiin kuvata. Mallin löytämisessä auttaa usein hajontakaavio, jonka avulla saadaan muodostettua käsitys siitä, millainen malli pistejoukkoon sopisi parhaiten. Tavoitteena on löytää sellainen matemaattinen malli, joka kuvaisi pistejoukon sijoittumista parhaiten. (Heikkilä, 2014, 222.)

Jos esimerkiksi pistejoukkoon voidaan luontevasti sovittaa suora, lineaarinen malli eli pienimmän neliösumman regressiosuora sopii hyvin muuttujien välisen yhteyden kuvaamiseen. Ei-lineaarisetkin riippuvuudet on usein mahdollista pakottaa muuttujamuutosten avulla lineaariseen malliin. (Heikkilä, 2014, 223.) Kaikista yksinkertaisimmassa lineaarisessa

mallissa on mukana vain yksi selittävä muuttuja (x) ja selitettävä muuttuja (y). Tämän suoran yhtälö on muodossa $y = a + b_x + \varepsilon$. Yhtälössä a on vakio, joka ilmaisee suoran ja y -akselin leikkauspisteen, b puolestaan tarkoittaa selittävän muuttujan regressiokerrointa, eli sitä kuinka paljon y -muuttuja keskimäärin muuttuu, kun x kasvaa yhden yksikön verran. Symboli ε merkitsee virhe- tai jäännöstermiä eli residuaalia. Residuaali kuvaa mallissa esiintyvää satunnaisvaihtelua eli virheen suuruutta, sillä usein malli ei kykene täysin selittämään ilmiötä, vaan mallissa on mukana virhettä tai ennustevajetta. Residuaalien tulee noudattaa mallissa normaalijakaumaa. (Heikkilä, 2014, 223; Metsämuuronen, 2011, 711 ja 713.) Residuaalia kutsutaan usein myös jäännöstermiksi eli jäännökseksi. Tässä tutkielmassa viitataan residuaaleihin usein jäännöksenä.

Regressioanalyysi menetelmänä voidaan jakaa kolmeen eri vaiheeseen, joista ensimmäisessä valitaan tarkasteltavat muuttujat. Näistä muuttujista ainakin yhden on oltava selitettävä ja muiden selittäviä. Toisessa vaiheessa tehdään varsinainen regressioanalyysi. Kolmannessa vaiheessa puolestaan arvioidaan saatuja tuloksia niin kutsuttujen ”diagnostisten tarkastelujen” avulla. (Metsämuuronen, 2011, 710–711.) Käytännössä viimeinen vaihe tarkoittaa analyysin tulosten luotettavuuden tarkastelua, esimerkiksi sitä, kuinka merkitsevästä mallin tuottamia tuloksia voidaan pitää.

Tässä tutkielmassa päädyttiin käyttämään lineaarista regressiomallia, sillä muuttujien väliset yhteydet olivat lineaarisia. Lisäksi tutkielmassa regressioanalyysia käytetään konfirmatorisesti, sillä tutkielmassa halutaan selvittää jo ennakolta tärkeiksi tiedettyjen muuttujien osuuksia selitettävässä ilmiössä. Toisaalta voidaan ajatella, että tutkielmassa tehdään myös eksploratiivista analyysia, sillä tutkielmassa seulotaan tiettyjä samoja ominaisuuksia kuvaavien muuttujien joukosta parhaiten selittävät. Kovin suureksi tätä seulomista ei voi kuitenkaan väittää, sillä tutkimusasetelmassa mukana on vain kahta ominaisuutta kuvaavia muuttujia. Yhteensä selittäviä muuttujia tutkimuksessa on viisi, joista kaksi kuvaa lähialue-estimaatin tulotasoa ja kolme puolestaan sen koulutustasoa. Näistä seulotaan parhaiten selitettävää muuttujaa – yläkoulujen päättökeskiarvoja – kuvaavat muuttujat.

Lineaarisen regressiomallin merkitsevyyttä tarkastellessa tulee ottaa huomioon se, ovatko mallille asetetut edeltävyys ehdot voimassa. Edeltävyys ehdot ovat seuraavanlaisia: Selitettävän ja selittävien muuttujien on oltava keskenään lineaarisessa yhteydessä. Muuttujien on myös oltava virheettömästi mitattuja ja niiden tulee olla intervalli- tai suhdeasteikollisia sekä jatkuvia. Lisäksi regressiomallin tuottamien jäännösten on oltava normaalisti jakautuneita ja

niiden summan on oltava lähellä nollaa. Lisäksi jäännösten oletetaan olevan *homoskedastisia*, eli mallin tuottamien jäännösten varianssin oletetaan olevan suunnilleen sama kaikilla muuttujien arvoilla. Lopuksi oletuksena on, että havainnot ovat toisistaan riippumattomia. (Jokivuori & Hietala, 2014; Osborne, 2020, 31–37, 56–58.; Valli, 2015.) Taanilan (2020a, 11) esittämänä lineaaristen regressiomallien edeltävyys ehdot ovat seuraavat:

- Selittävien ja selitettävän muuttujan välillä on lineaarinen riippuvuus.
- *”Jäännösten varianssi on yhtä suuri kaikilla selitettävien muuttujien arvoilla.*
- *Jäännökset noudattavat normaalijakaumaa*
- *Jäännökset ovat toisistaan riippumattomia.”*

Lineaarisuuden ja jäännösten varianssien yhtäsuuruuden tarkastelu voidaan tehdä jäännöskuvion avulla ja parhaiten tähän tarkasteluun sopii Taanilan (2020a, 12–13) mukaan hajontakuviio, jossa x-akselilla on esitetty mallin ennustamat arvo ja y-akselilla puolestaan vastaavat jäännökset. Samaa menetelmää suosittelee myös Osborne (2020, 56). Jos hajontakuvioista on mahdollista havaita johonkin suuntaan kaartuva havaintopisteiden muoto, tämä saattaa kertoa ei-lineaarisesta riippuvuudesta (Osborne, 2020, 56). Jos havaintopisteiden varianssi on suurin piirtein samankaltainen koko kuviossa, eivätkä pisteet muodosta esimerkiksi kartiota, voidaan varianssin yhtäsuuruuden ehtoa pitää toteutettuna (Osborne, 2020, 58).

Jäännösten normaalijakautuneisuutta puolestaan voidaan tutkia pienemmissä otoksissa silmämääräisesti normaalijakaumakuvioista tai histogrammista, sekä laskennallisesti käyttämällä apuna normaalijakaumaa testaavia testejä. Usein testeillä on kuitenkin taipumusta hylätä normaalijakaumaoletus liian helposti, minkä vuoksi usein suositellaankin käytettäväksi graafisia, silmämääräisiä menetelmiä. (Metsämuuronen, 2011, 610–611, 979; Osborne, 2020, 57; Taanila, 2020a, 14; Taanila, 2020c, 4–5, 8.) Taanilan suosittelemassa normaalijakaumakuviossa pysty- ja vaaka-akselit on laadittu muunnosten avulla sellaisiksi, että havainnot, jotka noudattavat normaalijakaumaa sijoittuvat suoralle viivalle kuviossa. Muuttujan havainnot voidaan olettaa normaalisti jakautuneiksi, jos kuviossa ei näy suuria poikkeuksia tältä vasemmasta alakulmasta oikeaan yläkulmaan kulkevalta viivalta. (Taanila, 2020a, 14.) Myös histogrammin käyttö on mahdollista (Osborne, 2020, 57).

Lisävarmuutta normaalijakautuneisuuden tulkintaan voidaan hakea suorittamalla esimerkiksi normaalijakaumaa testaava Kolmogorovin-Smirnovin testi. Testi soveltuu sellaisiin tilanteisiin, joissa ollaan kiinnostuneita muuttujan jakauman muodosta. Testissä vertaillaan

toisiinsa kahta jakaumaa, jotka voivat olla esimerkiksi havaintopisteiden jakauma ja normaalijakauma. Kyseinen testi tunnetaan myös nimellä Lilleforsin testi, jos vertailussa käytettävä toinen jakauma on nimenomaan normaalijakauma. Testin heikkoutena pidetään, että se hylkää nollassa nollahypoteesin (nollahypoteesi: jakauma on normaalisti jakautunut) ehkä liiankin herkästi kuten muutkin testit. (Metsämuuronen, 2011, 611, 979–980.) Analyysissä käytetty ohjelma SPSS kuitenkin mahdollistaa myös toisen normaalijakaumaa testaavan testin, Shapiro-Wilkin-testin tekemisen samaan aikaan, ja kummankin tuloksia yhtäaikaaisesti tarkastelemalla saadaan lisävarmuutta tuloksille. Näiden kahden testin p-arvot voivat poiketa toisistaan, mutta jos kummassakin testissä tullaan samaan päätelmään nollahypoteesista, ei testien eroja ole Taanilan (2020b, 8) mukaan syytä pohdiskella. Sen sijaan hän toteaa, että jos testien päätelmät ovat erilaiset, asiaa kannattaa tarkastella esimerkiksi histogrammin avulla.

Lineaariseen regressioanalyysiin malleja muodostettaessa havaittiin, että koulutusasteiden suhteellisten osuuksien väliset korrelaatiot olivat hyvin suuria, mikä aiheutti *multikolinearisuutta* malleissa. Useat monimuuttujamenetelmät ovat Metsämuuronen (2011, 645) mukaan analyysiin mukaan tulevien muuttujien voimakkaalle keskinäiselle korrelaatiolle. Multikolinearisuutta saattaa esiintyä esimerkiksi silloin, kun kaksi tunnuslukua mittaa lähes samaa asiaa. Multikolinearisuus aiheuttaa muun muassa regressioanalyysissä sen, että kaksi voimakkaasti toisiinsa yhteydessä olevaa muuttujaa tulee mukaan malliin, vaikka todellisuudessa vain toinen näistä lisää mallin selitystasetta, toisen ollessa turha. Sen lisäksi että liian voimakkaasti korreloivat muuttujat voivat olla turhia, ne saattavat aiheuttaa ongelmia myös tilastollisessa käsittelyssä ja ohjelmien sisäisessä työskentelyssä (Metsämuuronen, 2011, 645–646). Tämän vuoksi multikolinearisuudesta tulisi päästä eroon, jos tavoitteena on mahdollisimman luotettava lineaarinen regressioanalyysi.

Multikolinearisuuden vuoksi lisättäessä regressioanalyysissä muodostettaviin malleihin kaikkien eri koulutusasteiden osuudet alueilla, malleista putosi SPSS:n käsittelyssä pois korkeakoulutuksen saaneiden osuus. Tämä johtui siitä, että korrelaatio keskiasteen koulutuksen saaneiden osuuden ja korkeakoulutuksen saaneiden osuuden välillä oli erittäin voimakas ja tilastollisesti erittäin merkitsevä ($r_s = -,986$, $p < ,001$). Alla on esitetty ilman SIKU-alueiden rajoja muodostettujen lähialue-estimaattien sisältä poimittujen koulutustietojen keskinäiset korrelaatiot. SIKU-alueiden avulla muodostettujen estimaattien vastaavia korrelaatioita ei esitetä tässä, vaan tyydytään ainoastaan toteamaan niiden olleen käytännössä täysin samanlaisia.

Taulukko 3. Koulutusasteiden osuuksien keskinäiset korrelaatiot.

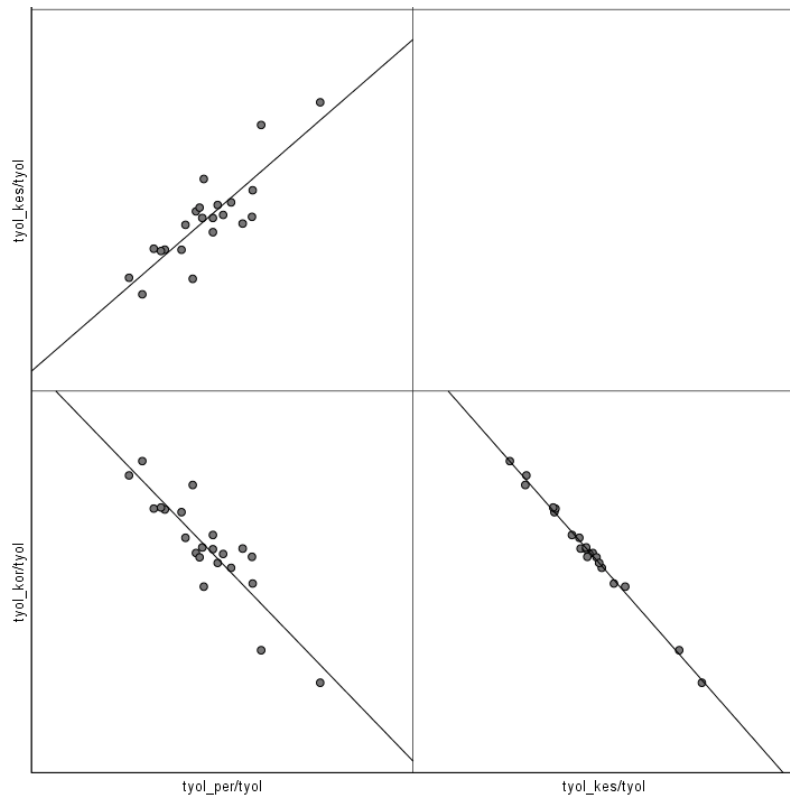
Correlations

			tyol_per/tyol	tyol_kes/tyol	tyol_kor/tyol
Spearman's rho	tyol_per/tyol	Correlation Coefficient	1,000	,802**	-,850**
		Sig. (2-tailed)	.	,000	,000
		N	22	22	22
	tyol_kes/tyol	Correlation Coefficient	,802**	1,000	-,986**
		Sig. (2-tailed)	,000	.	,000
		N	22	22	22
	tyol_kor/tyol	Correlation Coefficient	-,850**	-,986**	1,000
		Sig. (2-tailed)	,000	,000	.
		N	22	22	22

** . Correlation is significant at the 0.01 level (2-tailed).

Voidaan todeta keskiasteen koulutuksen ja korkeakoulutuksen saaneiden osuuksien välillä vallitsevan täydellisyyttä lähentelevän negatiivisen riippuvuussuhteen ($r_s = -,986$, Sig. < ,01). Riippuvuussuhde voidaan myös todeta lineaariseksi ja erittäin voimakkaaksi hajontakaavion avulla. Näin ollen näiden kahden muuttujan välillä todellakin esiintyy multikolinearisuutta ja toinen niistä on syytä jättää pois regressiomallia muodostettaessa.

Kuvaaja 1. Koulutusmuuttujien keskinäiset yhteydet hajontakuviona



Korrelaatio ei sen sijaan ollut aivan yhtä voimakasta ($r_s = -,850$) perusasteen ja korkeakoulutuksen suhteellisten osuuksien välillä, mutta se kuitenkin erittäin korkeaa ja tilastollisesti erittäin merkitsevällä tasolla (Sig. < ,01). Lineaarisuus ja negatiivinen voimakas riippuvuus voidaan todeta myös hajontakaaviosta. Erittäin voimakas korrelaatio löytyi myös perusasteen ja keskiasteen osuuksien väliltä ($r_s = ,802$), mutta korrelaatio oli tässä tapauksessa positiivinen; toisin sanottuna perusasteen koulutuksen osuuden lisääntyessä myös keskiasteen koulutuksen osuus lisääntyy työssäkäyvän työväestön keskuudessa. Lisäksi tämäkin riippuvuussuhde oli aineistossa erittäin merkitsevällä tasolla (Sig. < ,01). Voimakas lineaarinen riippuvuus muuttujien välillä voidaan todentaa myös hajontakuviota tarkastelemalla.

4 Analyysi ja tulokset

Tässä luvussa esitellään tutkimustulokset ja kuvataan niihin johtaneet prosessit. Tuloksien merkitystä ei tässä luvussa pohdita tarkemmin, vaan se tapahtuu seuraavassa pääluvussa. Tämän luvun pääasiallisena tarkoituksena onkin esittää lukijalle se prosessi, jonka seurauksena tuloksiin on päädytty, sekä tietenkin kuvata tulokset. Tämä luku on jaettu kahteen alalukuun siten, että ne esitetään kahden vaihtoehtoisen lähialue-estimaattien muodostamisen skenaarion perusteella. Lisäksi nämä alaluvut on jaettu omiksi alaluvuikseen siten, että kummankin lähialue-estimaattiskenaarion kohdalla omissa luvuissaan esitellään työväestön koulutuksen, sekä lähialueen tulojen vaikutukset oppimistuloksiin yläkouluissa.

Lukijan kannattaa huomata, että korrelaatioanalyysi on tehty mahdollisesti kahdella eri tavalla: toisessa tavassa poikkeavat havainnot ovat korrelaation tutkimisessa mukana, toisessa tavassa puolestaan eivät. Lisäksi regressioanalyysien tulokset on esitelty kummankin lähialue-estimaattiskenaarion kohdalla omassa alaluvussaan. Myös regressioanalyysien kohdalla kannattaa huomata, että kummassakin skenaariossa on muodostettu kaksi lineaarista regressiomallia, jotka esitetään omien väliotsikoidensa alla.

Jotta tuloksia olisi helpompi lukea ja ymmärtää niiden takana vaikuttavia prosesseja, havainnollistetaan tuloksia ja niihin johtaneita vaiheita lukuisten taulukoiden ja erilaisten kuvaajien avulla. Aluksi tuloksiin johtavien prosessien kuvaus on melko yksityiskohtaista, mutta jatkossa samanlaisia prosesseja ei enää kuvata samalla tarkkuudella, vaan asiat pyritään lähinnä toteamaan ja etenemään nopeasti varsinaiseen tulosten kuvaamiseen.

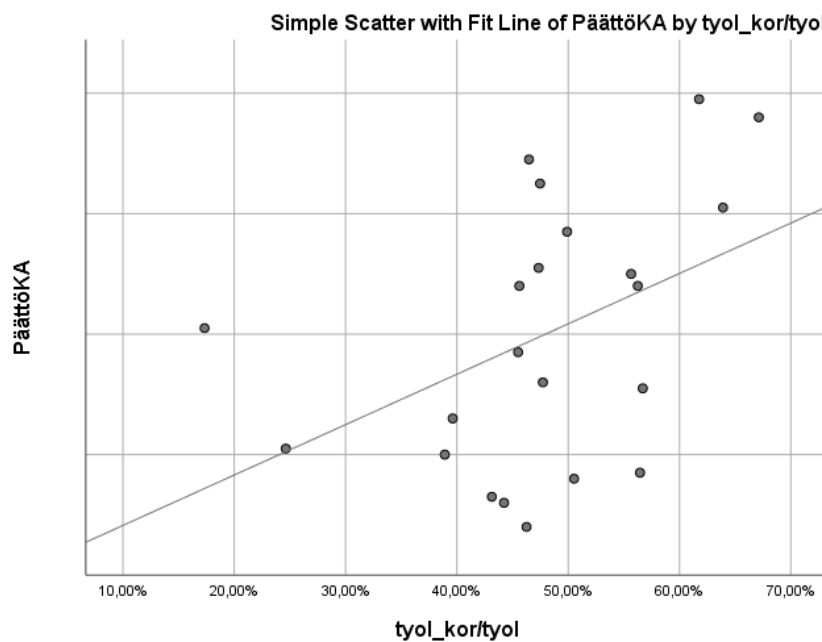
4.1 Yläkoulujen lähialue-estimaatit ilman oppilasalueita

Tässä alaluvussa esitellään tutkimuksessa saatuja tuloksia sosioekonomisten tekijöiden yhteyksistä yläkoulujen oppimistuloksiin kuvitteellisessa skenaariossa, jossa Oulun kaupungilla ei olisi lainkaan oppilasalueina toimivien sivistys- ja kulttuurialueiden (SIKU-alueet) rajoja käytössä oppilaille lähikouluja määriteltäessä. Ensimmäisenä esitellään tässä skenaariossa muodostettujen lähialue-estimaattien työssäkäyvän työväestön koulutustason yhteyksiä oppimistuloksiin. Toisessa alaluvussa puolestaan käsitellään samaisessa skenaariossa lähialue-estimaatin asukkaiden laskennallisten tulojen yhteyksiä oppimistuloksiin. Kolmannessa alaluvussa esitellään lineaarisen regressioanalyysin tulokset kootusti.

4.1.1 Työväestön koulutus ja oppimistulokset

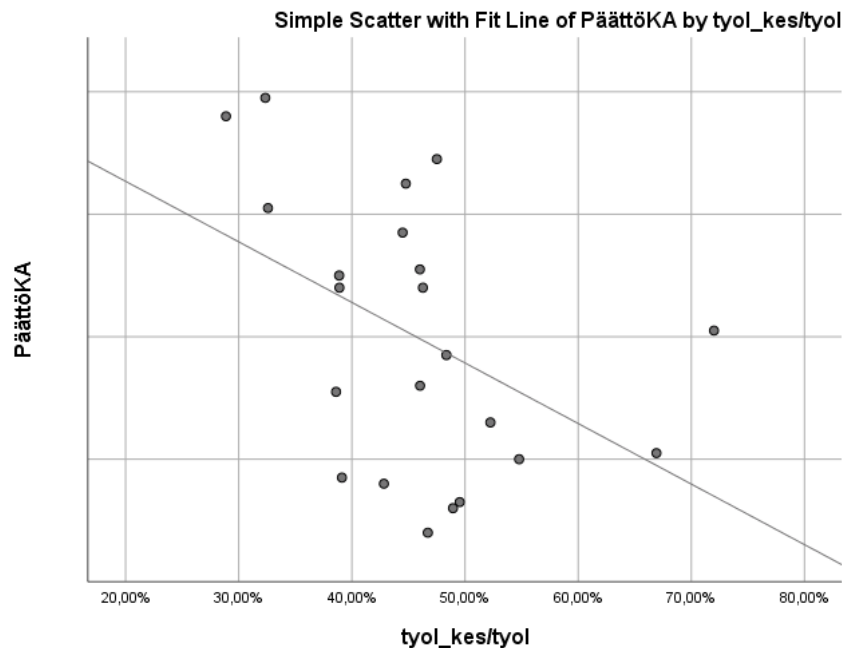
Työväestön koulutusasteen ja päättökeskiarvojen välistä yhteyttä tarkasteltiin aluksi hajontakuvioiden kautta. Hajontakuvioiden avulla pystytään havaitsemaan, että työssäkäyvän työväestön korkeakoulutuksella on positiivinen korrelaatio yläkoulujen päättökeskiarvoihin. Kuviota tutkittaessa voidaan huomata, että havaintopisteet ovat jakautuneet jokseenkin selkeästi muotoon, joka voidaan tulkita lineaariseksi. Hajontakuviosta kuitenkin huomataan, että aineistossa on kaksi melko selkeää *outlieria*, eli poikkeavaa havaintoa, jotka muuttavat regressiosuoran kulmaa loivemmaksi. Jos nämä kaksi poikkeavaa havaintoa poistettaisiin aineistosta, olisi regressiosuoran kulma todennäköisesti jyrkempi ja mallin selitysaste (näin ollen mahdollisesti suurempi).

Kuvaaja 2. Päättökeskiarvon ja korkeakoulutetun työväestön osuuden hajontakaavio.



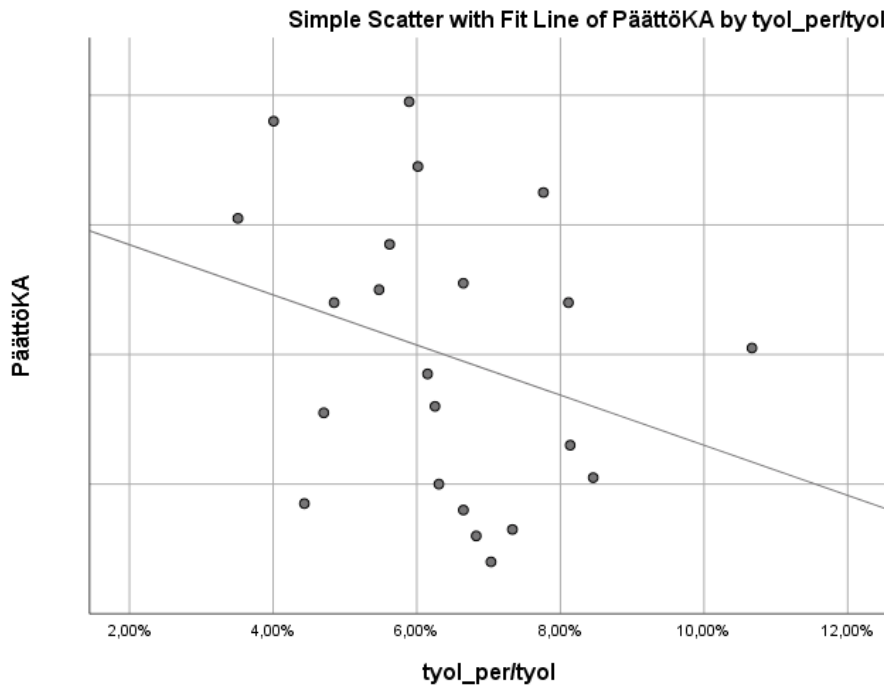
Keskiasteen koulutuksen saaneen työväestön ja päättökeskiarvojen väliseen hajontakuviioon piirretty regressiokuvaaja on puolestaan laskeva, joten näiden välillä on negatiivinen korrelaatio: keskiasteen koulutuksen saaneiden määrän kasvaessa päättökeskiarvot laskevat. Keskiasteen koulutuksen suhteellisen osuuden ja päättökeskiarvon välisestä hajontakuviosta voidaan tehdä myös havainto, että kyseessä on lähes sama kuvio kuin korkeakoulutettujen ja päättökeskiarvon välillä. Kuvio on käytännössä ainoastaan käännetty y-akselinsa suhteen peilikuvaksi. Näin ollen voidaan päätellä, että keskiasteen- ja korkeakoulutuksen saaneiden suhteelliset osuudet ovat toisistaan hyvin voimakkaasti riippuvaisia.

Kuvaaja 3. Päätökeskiarvon ja keskiasteen koulutuksen saaneen työväestön osuuden hajontakaavio.



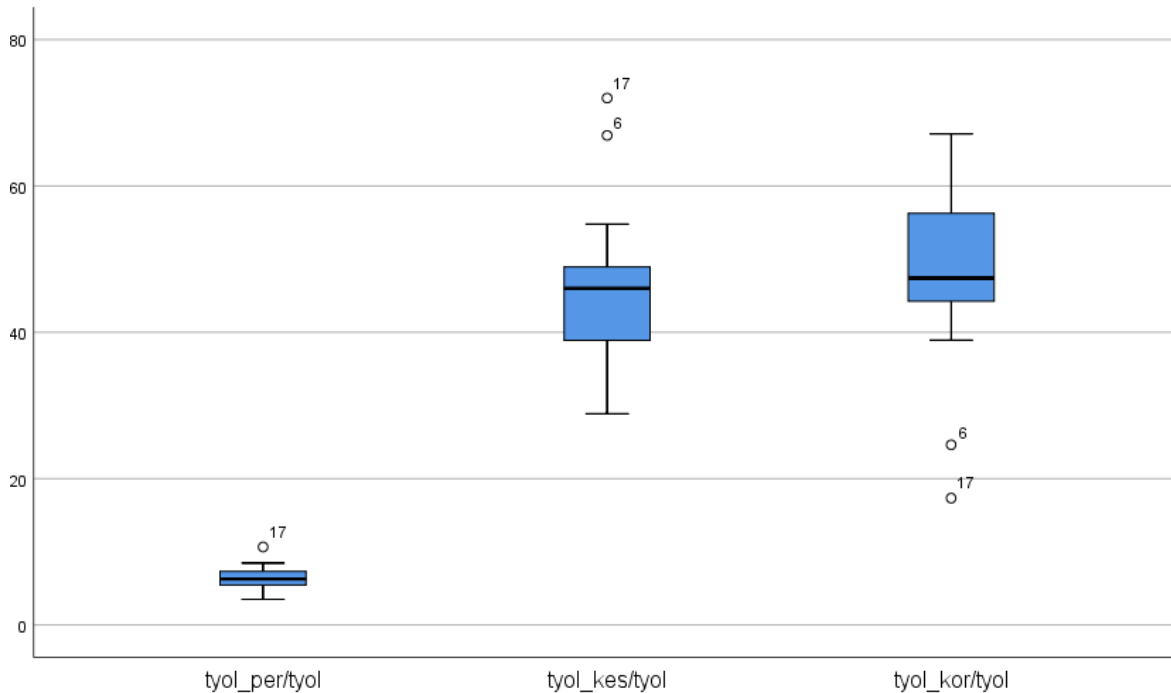
Pelkästään perusasteen koulutuksen saaneen työssäkäyvän työvoiman osuuden kohdalla voidaan huomata sama asia kuin keskiasteen koulutuksen kohdalla: perusasteen koulutuksen saaneen työväestön suhteellisen osuuden kasvaessa yläkoulujen oppimistulokset heikkenevät päätökeskiarvoilla mitattuna. Havaintopisteet ovat kuitenkin levittäytyneet kuvaajassa laajemmin kuin muiden koulutusasteiden kohdalla vastaavassa silmämääräisessä tarkastelussa. Kuviosta havaitaan, että regressiosuora on loivemmin laskeva ja perusasteen koulutuksen saaneiden osuuden yhteys oppimistuloksiin vaikuttaisi näin ollen heikommin negatiiviselta kuin keskiasteen koulutuksen saaneen työvoiman yhteys. Myös tässä havaintokaaviossa on nähtävissä mahdollinen poikkeava havainto kuvion oikeassa reunassa.

Kuvaaja 4. Päätökeskiarvon ja perusasteen koulutuksen saaneen työväestön osuuden hajontakaavio.



Koska hajontakaavioiden perusteella pystyttiin epäilemään, että aineistossa on mukana poikkeavia havaintoja, oli asiaa syytä tutkia tarkemmin. Kaikki epäillyt havaintopisteet voitiin varmistaa outliereiksi laatikko-janakuvion avulla. Laatikko-janakuvioista nähdään myös, että poikkeava havainto on saman koulun alue (havainto 17) kaikissa muuttujissa. Keski- ja korkeasteenkoulutusten kohdalla myös havainto 6 on sama. Tämän tiedon perusteella oli syytä pohtia poikkeavien havaintojen poistamista aineistosta. Tutkimuksessa päädyttiin havainnon myötä ratkaisuun, että analyysi toteutetaan sekä outliereiden kanssa, että ilman niitä. Perusteluna tälle on, että poikkeavat havainnot johtuvat aineiston keräämiseen ja käsittelyyn liittyvistä tekijöistä. Jos poikkeavat havainnot jätetään kokonaan käsittelemättä, ei saada joiltakin alueilta lainkaan tietoa. Toisaalta näiden havaintojen mukana oleminen vaikutti aineistoon siten, että se korosti tietynlaisten asuinalueiden ominaisuuksia vääristäen vallitsevaa todellisuutta. Poikkeavien havaintojen plussien ja miinusten mennessä harkinnassa vaakakuppeihin melko tasan, päädyttiin toteuttamaan analyysi kahdella tavalla. Poikkeavien havaintojen ollessa mukana täytyy kuitenkin osoittaa kriittisyyttä analyysiä kohtaan, sillä näiden havaintojen kohdalla poimittujen tietojen määrä on jopa alle 5 % alueen todellisesta tietomäärästä.

Kuvaaja 5. Työllisen työväestön koulutustasojen osuuksien laatikko-janakuvio.



Analysoitaessa havaintoja poistamatta edellä huomattuja poikkeavia havaintoja huomattiin, että Pearsonin korrelaatiokertoimen kaikki ehdot eivät aineistossa täyty juuri havaittujen outlier-tapauksien vuoksi. Lisäksi muuttujien normaalijakautuneisuudesta ei pystytty olemaan täysin varmoja, vaikka normaalisuutta testaavien testien mukaan kaikki muuttujat olivatkin normaalisti jakautuneita. Muuttujien normaalijakautuneisuutta histogrammien ja normaalijakaumakuvioiden tarkastelujen perusteella kuitenkin päädyttiin siihen, että normaalijakauman toteutumisesta ei voida aineiston pienehkön koon vuoksi olla täysin varmoja. Tämän vuoksi päädyttiin korrelaatioanalyysissä käyttämään poikkeavien havaintojen mukana ollessa Spearmanin järjestyskorrelaatiokerrointa. Spearmanin järjestyskorrelaatiokerrointa on suositeltavaa käyttää Pearsonin korrelaatiokertoimen sijaan aineistoissa, joissa on poikkeavia havaintoja, sillä se ei reagoi yhtä herkästi poikkeaviin muuttujiin. Spearmanin järjestyskorrelaatiokerroin ei myöskään vaadi muuttujien normaalijakaumaa tai sitä, että muuttujat ovat jatkuvia. Lisäksi Spearmanin järjestyskorrelaatiokerroin toimii myös silloin kun aineiston yhteydet eivät ole täysin lineaarisia, sillä Spearmanin järjestyskorrelaatiossa monotoninen yhteys muutetaan lineaariseksi. (Schober ym., 2018, 1765–1766; Valli, 2015.) Spearmanin järjestyskorrelaatio ei sen sijaan perinteisessä muodossaan tuota oikeaa tulosta, jos aineistossa esiintyy sidoksia eli samoja arvoja kummassa tahansa mukana olevassa muuttujassa (Metsämuuronen, 2011, 367).

Tutkimuksen aineistossa missään muuttujassa ei kuitenkaan esiintynyt riippuvuuksia, joten Spearmanin järjestyskorrelaatiokerrointa voidaan hyvin käyttää.

Analyysissä korkeimman Spearmanin järjestyskorrelaatiokertoimen arvon ($r_s = -,515$) sai keskiasteen koulutuksen saaneen työssäkäyvän työväestön suhteellinen osuus. Tätä voidaan kuvailla ”melko korkeaksi” tai ”kohtuulliseksi” (engl. ”moderate”) korrelaatioksi (Akoglu, 2018, 92; Metsämuuronen, 2011, 371). Korrelaatiota voidaan pitää myös tilastollisesti merkitsevä (Sig. = ,014). Korkeakoulutuksen kohdalla korrelaatiota ($r_s = ,486$) voidaan myös pitää melko korkeana, ja myös tämän muuttujan kohdalla korrelaatiota voidaan pitää tilastollisesti merkitsevä (Sig. = ,022). Sen sijaan pelkän peruskoulun käyneiden osuuden kohdalla korrelaatiota ($r_s = -,388$) voisi pitää ”heikon” ja ”kohtuullisen” välimaastoon osuvana määritellyistä rajoista riippuen, eikä sitä voida pitää myöskään tilastollisesti merkitsevä (Sig. = ,074).

Taulukko 4. Työssäkäyvän työväestön koulutusasteen ja päättökeskiarvojen korrelaatio.

Correlations						
			PäättöK A	tyol_per/ty ol	tyol_kes/ty ol	tyol_kor/ty ol
Spearman's rho	PäättöKA	Correlation Coefficient	1,000	-,388	-,515*	,486*
		Sig. (2-tailed)	.	,074	,014	,022
		N	22	22	22	22

*. Correlation is significant at the 0.05 level (2-tailed).

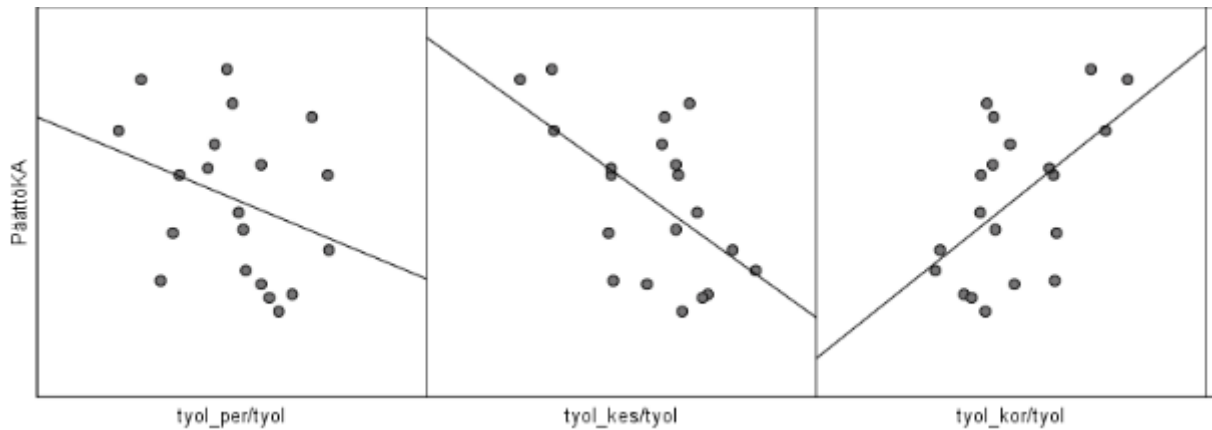
Työväestön koulutus ja päättökeskiarvot ilman poikkeavia havaintoja

Kun aineistosta poistettiin poikkeavat havaintopisteet 6 ja 17, voitiin tarkastella uudestaan Pearsonin korrelaatiokertoimen ennakkoehtoja. Muuttujien asteikot eivät luonnollisesti muuttuneet, joten tältä osin ehdot ovat voimassa.

Muuttujien lineaarinen yhteys voidaan todeta lyhyesti hajontakuviomatriisista. Näyttäisi siltä, että muuttujista perusasteen koulutuksen saaneiden osuuden yhteys päättökeskiarvoon ei ole kovinkaan vahva. Laskeva regressiosuora voidaan kuitenkin piirtää ja lineaarisuus todeta.

Toisten muuttujien yhteys selitettävään muuttujaan on sen sijaan selkeämpi. Todetaan että lineaarisuusehto on voimassa.

Kuvaaja 6. Päätöskeskisarvon ja koulutusmuuttujien väliset hajontakuviot.



Muuttujien normaalijakautuneisuutta voidaan tarkastella sekä normaalijakaumakuvion että histogrammien kautta, tai normaalijakautuneisuutta testaavien testien avulla. Tutkimuksessa hyödynnettiin näitä kaikkia. Tässä yhteydessä kuitenkin esitetään tilan säästämiseksi vain testien tulokset ja tyydytään toteamaan, että kuvaajista on nähtävissä samat tulokset. Normaalijakaumaa testaavien Kolmogorov-Smirnovin ja Shapiro-Wilkin testien perusteella voitiin todeta, että kaikkien muuttujien kohdalla nollahypoteesi jää voimaan ($\text{Sig.} > 0,05$) ja näin ollen kaikkia muuttujia voidaan pitää normaalisti jakautuneina. Histogrammien ja normaalijakaumakuvioiden avulla tehty silmämääräinen tarkastelu kuitenkin herätti epäilyksiä aineiston normaalijakautuneisuudesta lähes kaikkien muuttujien kohdalla ja näin ollen päädyttiin luopumaan kokonaan Pearsonin korrelaatiokertoimen käyttämisestä ja jatkamaan Spearmanin järjestyskorrelaatiokertoimen käyttöä. Tämä mahdollistaa myös tulosten paremman vertailun keskenään.

Taulukko 5. Normaalijakauman testit työllisyyttä kuvaaville muuttujille.

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
PäättöKA	,125	20	,200*	,942	20	,265
tyol_per/tyol	,092	20	,200*	,973	20	,808
tyol_kes/tyol	,164	20	,166	,949	20	,355
tyol_kor/tyol	,189	20	,058	,937	20	,210

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Saadut Spearmanin järjestyskorrelaatiokertoimen arvot osoittavat, että kaikista voimakkain yhteys päättökeskiarvoon koulutusmuuttujilla, joiden joukosta oli poistettu poikkeavat havainnot, oli keskiasteen koulutuksen saaneen työväestön suhteellisella osuudella. Yhteys oli kohtuullisen voimakasta, negatiivista ja tilastollisesti merkitsevällä tasolla ($r_s = -,547$, Sig. = ,013). Korkeakoulutettujen osuudella oli tilastollisesti merkitsevä ja kohtuullisen voimakas yhteys yläkoulujen päättökeskiarvoon ($r_s = -,517$, Sig. = ,020). Sen sijaan pelkän peruskoulun käyneiden työssä olevien osuudella lähialue-estimaatin alueella asuvasta työvoimasta ei ollut tilastollisesti merkitsevää yhteyttä, mutta yhteys oli kuitenkin kohtuullisen voimakkasta ja negatiivista ($r_s = -,408$, Sig. = ,074).

Taulukko 6. Päätökeskiarvojen ja koulutusmuuttujien väliset korrelaatiot.

Correlations						
			PäätöKA	tyol_per/tyol	tyol_kes/tyol	tyol_kor/tyol
Spearman's rho	PäätöKA	Correlation Coefficient	1,000	-,408	-,547*	,517*
		Sig. (2-tailed)	.	,074	,013	,020
		N	20	20	20	20

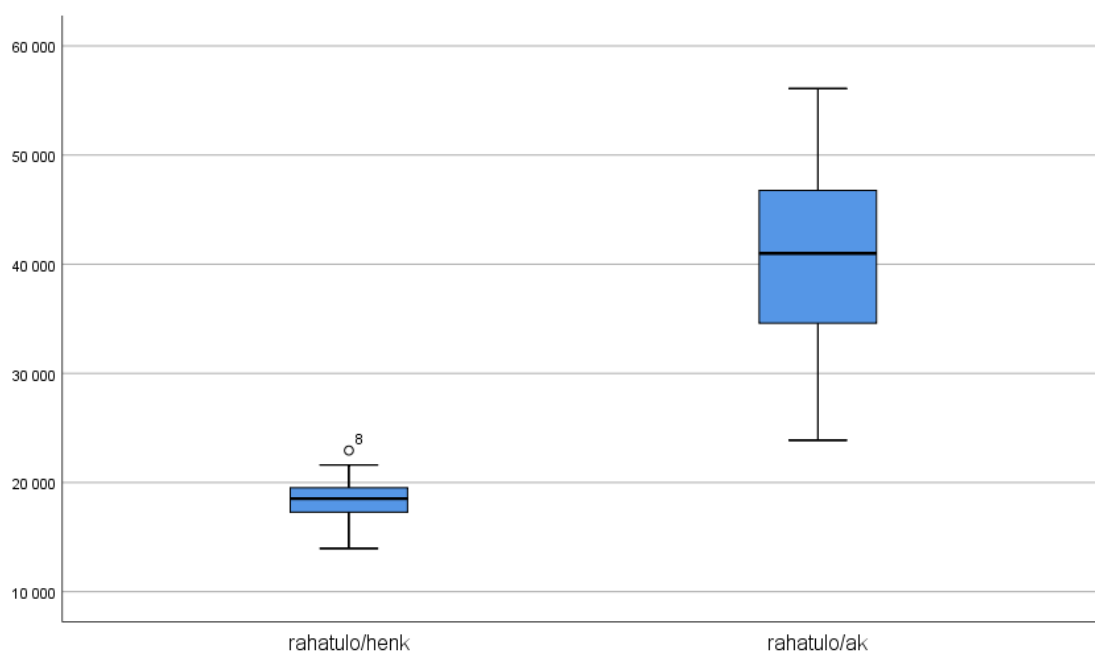
*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

4.1.2 Lähialueen tulot ja oppimistulokset

Yläkoulujen ilman oppilasalueiden rajoja muodostettujen lähialue-estimaattien tulojen yhteyksiä päätökeskiarvoihin analysoitaessa päädyttiin tutkimaan tuloja kahdella erilaisella tavalla. Ensimmäinen tapa oli laskea lähialue-estimaatin sisällä keskimääräinen *rahatulo per henkilö* (rahatulo/henk). Toinen tapa puolestaan oli laskea alueen rahatuloista asuntokuntakohtainen keskiarvo eli *rahatulo per asuntokunta* (rahatulo/ak). Näiden kahden muuttujan välisiä yhteyksiä suhteessa lähialue-estimaatin päätökeskiarvoon tutkittiin Spearmanin järjestyskorrelaatiokertoimen avulla, sillä Pearsonin tulomomenttikorrelaatiokertoimen ehdoista muuttujien normaalijakautuneisuuden ehtoa ei pystytty aukottomasti toteamaan voimassa olevaksi graafisten silmämääräisten tarkastelujen tai normaalijakautuneisuutta testaavien testien avulla. Testien perusteella muuttujat olisivat normaalijakautuneita, mutta aineiston suhteellisen pienen koon takia silmämääräiset tarkastelut normaalijakaumakuviosta tai muuttujien jakaumaa kuvaavista histogrammeista herättivät epäilyksiä todellisesta normaalijakautuneisuudesta. Lisäksi laatikko-janakuvista kävi ilmi, että muuttujien joukossa oli yksi poikkeuksellisen korkea havainto tulot per henkilö -muuttujan kohdalla. Aineiston tarkempi tutkiminen ei antanut syytä olettaa, että kyseessä olisi virhe, vaan korkea arvo johtuu todennäköisesti kyseisen lähialue-estimaatin oikeista ominaisuuksista. Tutkittaessa mikä alue on kyseessä, huomattiin nimittäin, että alueella todellakin asuu keskimääräistä parempituloisia ihmisiä. Edellä kuvattujen tekijöiden vuoksi varmuuden vuoksi päädyttiin käyttämään Spearmanin järjestyskorrelaatiokerrointa, jota normaalijakautuneisuuden mahdollinen puuttuminen tai aineistossa olevat poikkeavat havainnot eivät häiritse yhtä voimakkaasti kuin Pearsonin tulomomenttikorrelaatiota.

Kuvaaja 7. Tulomuuttujien laatikko-janakuviot.



Analyysissä havaittiin, että Spearmanin järjestyskorrelaatiokertoimella mitattuna löytyy heikko yhteys päättökeskiarvon ja rahatulon per henkilö väliltä ($r_s = ,377$). Asuntokuntaakohtaisen tulon kohdalla korrelaatio päättökeskiarvon kanssa oli myös heikko ($r_s = ,245$). Kummankaan muuttujan kohdalla p-arvo (Sig. = ,084 ja Sig. = ,272) ei osoittanut tilastollista merkitsevyyttä.

Taulukko 7. Laskennallisten rahatulokeskiarvojen suhde päättökeskiarvoon.

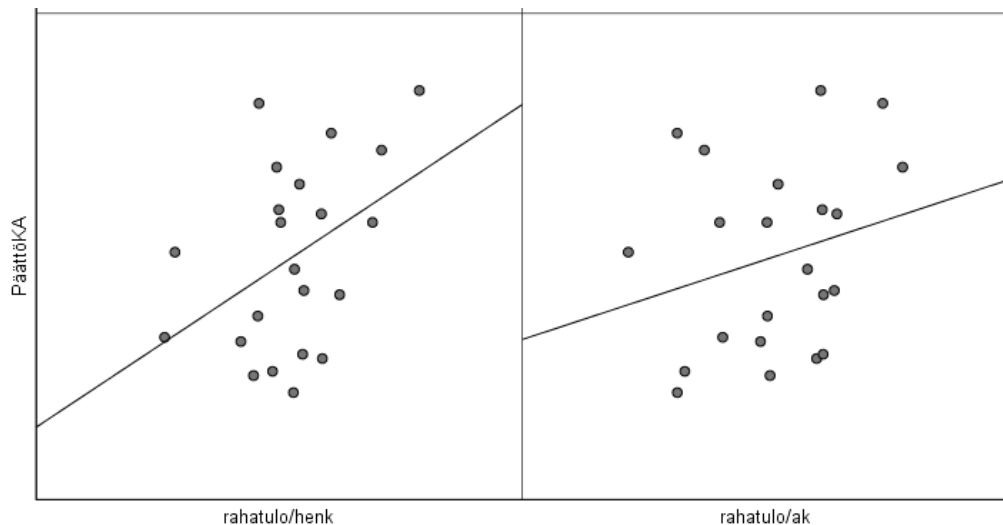
Correlations

			PäättöKA	rahatulo/henk	rahatulo/ ak
Spearman's rho	PäättöKA	Correlation Coefficient	1,000	,377	,245
		Sig. (2-tailed)	.	,084	,272
		N	22	22	22

Yläkoulujen päättökeskiarvojen ja laskennallisen rahatulon per henkilö hajontakuvaajasta voidaan havaita heikko lineaarinen riippuvuus. Jonkinlaisia poikkeavia havaintojakin kuvaajan perusteella aineistossa saattaisi olla, mutta tapauksien hajaantuneisuuden vuoksi tätä on vaikea silmämääräisesti varmistaa. Asia todettiin jo aiemmin laatikko-janakuvion avulla ja näin tiedetään aineistossa olevan mukana poikkeava havainto. Myös asuntokuntaakohtaisen

keskiarvoisen tulon ja päättökeskiarvojen välisen hajontakaavion avulla voidaan havaita heikko positiivinen korrelaatio. Kummankin näistä havainnoista osoitti jo varsinainen korrelaatioanalyysi.

Kuvaaja 8. Päättökeskiarvojen ja tulomuuttujien väliset korrelaatiot hajontakuvioissa.



Työväestön tulot ja päättökeskiarvot ilman poikkeavia havaintoja

Kun aiemmin havaittu poikkeuksellisen korkean arvon saanut havainto poistettiin aineistosta, havaittiin laatikko-janakuvioiden avulla, että aineistoon ilmestyi nyt toinenkin poikkeava havainto. Tällä kertaa poikkeava havainto kuvasi poikkeuksellisen matalia tuloja per henkilö. Poikkeava havainto päätettiin poistaa, sillä havainnon luotettavuutta ei voitu datan tarkastelun perusteella pitää hyvänä johtuen kyseisen lähialue-estimaatin vähäisistä hyödynnetyistä tiedoista (alle 5 % alueen tiedoista olivat hyödynnettäviä). Kun tämäkin havainto oli poistettu tarkastelusta, samaan muuttujaan ilmestyi vielä kolmas poikkeava havainto, joka päätettiin poistaa samoin perustein kuin edellinen poikkeava havainto. Näin lopulliseksi havaintojen määräksi (N) jäi 19 yläkoulujen lähialue-estimaattia. Koska tuloissa per asuntokunta ei alun perinkään ollut mukana poikkeavia havaintoja, eikä niitä ilmaantunut poistettaessa havaintopisteitä toisen muuttujan perusteella, ei tälle tuloille per asuntokunta suoritettu uutta korrelaatioanalyysiä.

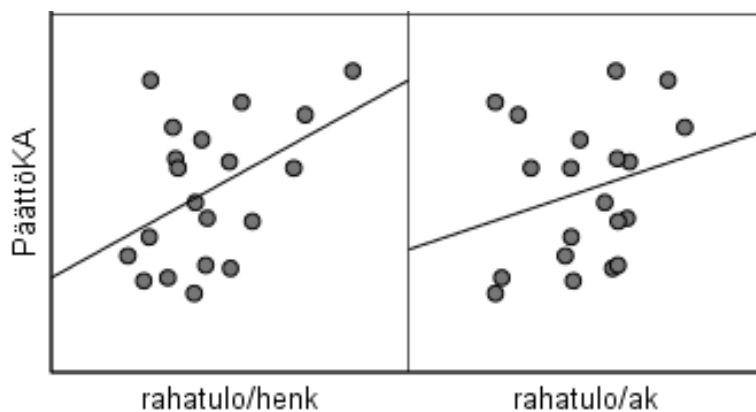
Uudessa korrelaatioanalyysissä päättökeskiarvojen ja rahatulojen per henkilö väliseksi yhteydeksi saatiin $r_s = ,272$ eikä yhteys ollut myöskään tilastollisesti merkitsevällä tasolla (Sig. = ,260). Otoskoko (N) oli 19. Voidaan todeta, että rahatulojen per henkilö yhteys

päätökeskiarvoihin itseasiassa heikkeni poikkeavien havaintojen poistamisen myötä. Sama asia voidaan havaita myös hajontakuviosta, joista lineaarisuus on regressiosuoran avulla enää heikosti nähtävissä.

Taulukko 8. Päätökeskiarvojen ja rahatulomuuttujien korrelaatio ilman poikkeavia havaintoja.

Correlations			PäätöKA	rahatulo/henk
Spearman's rho	PäätöKA	Correlation Coefficient	1,000	,272
		Sig. (2-tailed)	.	,260
		N	19	19

Kuvaaja 9. Päätökeskiarvon ja rahatulomuuttujien väliset hajontakuviot ilman poikkeavia havaintoja.



4.1.3 Regressioanalyysien tulokset

Regressioanalyysia suoritettaessa todettiin, että mallin luotettavuudesta saadaan korkeampi, mikäli poikkeavia havaintoja koulutustasoja kuvaavissa muuttujissa sisältävät havainnot poistetaan kokonaan aineistosta ennen regressiomallien luomista. Näiden havaintojen poistaminen kokonaan perustuu aineiston tuntemiseen ja siihen, että kyseisten havaintojen kohdalla sekä koulutus-, että tulotasoa sisältäviä tilastotietoruutuja pystyttiin poimimaan tietosuojatekijöiden vuoksi poikkeuksellisen vähän määrän jäädessä kummassakin havaintopaikassa jopa alle 5 % kyseisillä lähialue-estimaateilla olevien ruutujen määrästä. Näin voidaan todeta, että kyseisten havaintojen luotettavuus on erittäin heikko, eikä niitä ole syytä

pitää mukana regressioanalyysissä, jos halutaan mahdollisimman hyvin kuvaavia malleja. Samalla poistuivat poikkeuksellisen korkeiden tulojen ”piilottamat” poikkeuksellisen matalat tulot tarkastelusta, sillä ne liittyivät samoihin havaintopaikkoihin kuin koulutustasoa kuvaavat poikkeukselliset havainnot.

Sen sijaan tuloja kuvaavien muuttujien osalta päädyttiin jättämään poikkeuksellisen korkeat tulot sisältävä havainto mukaan aineistoon. Syynä menettelyyn oli, että poikkeuksellisen korkeat tulot sisältävän havainnon voidaan katsoa olevan luotettava aineiston tuntemisen perusteella. Kyseiseltä alueelta ruuduista on saatu käyttöön merkittävä osa, eikä virheitä ole syytä epäillä. Näin ollen myös tämän poikkeavan arvon voi katsoa kuuluvan aineistoon sellaisenaan. Lukijan kannattaa huomata, että myös tuloja kuvaavien muuttujien osalta on poistettu täsmälleen samat havaintopisteet kuin koulutusta kuvaavien muuttujien osalta. Ennakkoon tiedettiin näiden havaintojen olevan poikkeavia, jos suuret tulot sisältävä havainto poistetaan. Näin ollen aineiston tuntemisen perusteella, vedoten kyseisten kahden havainnon erittäin heikkoon luotettavuuteen päätettiin kyseiset havainnot jättää regressioanalyyseistä täysin pois.

Ensimmäinen malli

Ensimmäisenä haluttiin muodostaa malli, joka selittää parhaiten yläkoulujen päättökeskiarvojen vaihtelua. Mallin löytämiseksi käytettiin SPSS:n lineaarisen regressioanalyysin Stepwise-algoritmia. Stepwise-algoritmissa SPSS valitsee malliin ensimmäisenä sen selittävän muuttujan, joka selittää selitettävän muuttujan vaihtelua parhaiten. Tämän jälkeen algoritmi lisää malliin aina sen muuttujan jäljellä olevista selittävästä, joka parhaiten selittää muuttujan vaihtelua. Mahdollisen multikolinearisuuden tai kolinearisuuden vuoksi jo aiemmin malliin otettujen muuttujien selitysvoima voi kuitenkin myös laskea usein muuttujien mukaan tulemisen vuoksi. Algoritmi voi tämän vuoksi myös pudottaa muuttujia mallista. Kun algoritmi ei löydä enää muuttujaa, joka lisää mallin selitysvoimaa riittävästi malli on valmis (Jokivuori & Hietala, 2014; Taanila, 2020a, 23.)

Ensimmäinen malli sisältää vain yhden muuttujan, sillä kaikista parhaiten selittäväksi malliksi nousi Stepwise-algoritmin avulla malli, jossa mukana on ainoastaan keskiasteen koulutuksen saaneen työllisen työväestön suhteellinen osuus. Toisin sanottuna lisättäessä malliin mitä tahansa muita muuttujia mallin selitysaste laski. Taulukosta havaitaan, että tarkistettu selityskerroin (Adjusted R^2) on mallilla 0,297. Voidaan siis todeta, että malli selittää 29,7 %

yläkoulujen päättökeskiarvojen vaihtelusta. Sivuhuomiona mainittakoon, että korkeakoulutuksen saaneiden ollessa ainoana selittävänä muuttujana ei mallin selitysaste ($\text{Adjusted } R^2 = ,270$) poikennut kovin paljon tässä esitetystä mallista. Nämä kaksi koulutusastetta ovat kuitenkin multikolineaarisia keskenään ja keskiasteen koulutuksen osuuden selitysaste oli hieman parempi, minkä vuoksi juuri kyseinen malli esitetään. Mallissa mukana oli 20 yläkoulua lähialue-estimaatteineen, sillä kaksi poikkeavaa havaintopistettä poistettiin ennen mallin muodostamista.

Taulukko 9. Ensimmäisen mallin yhteenveto.

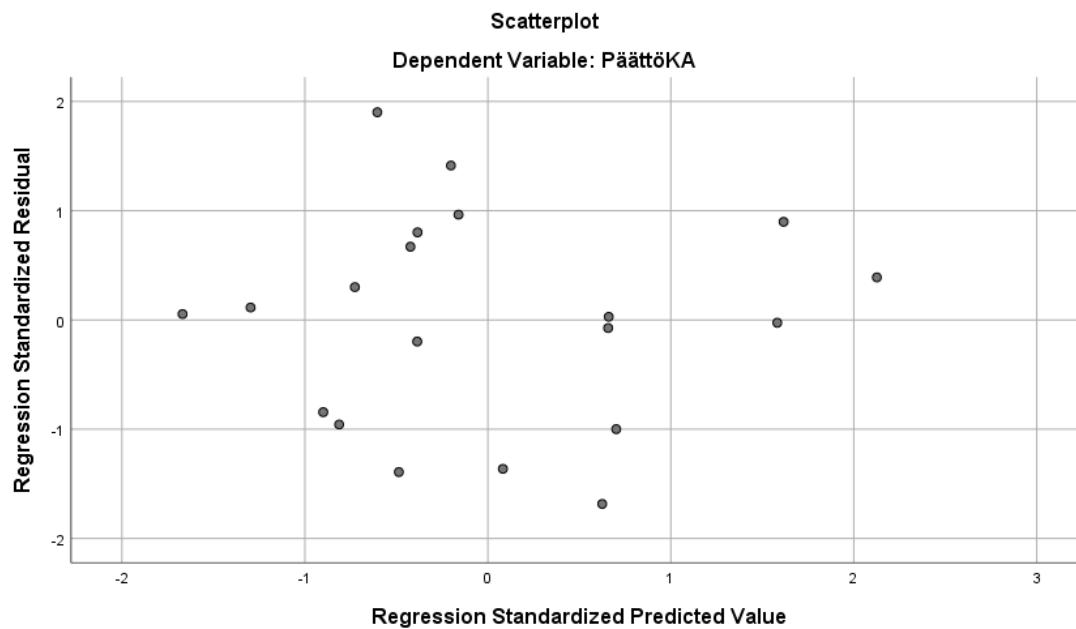
Model Summary ^b				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,578 ^a	,334	,297	,18990

a. Predictors: (Constant), tyol_kes/tyol

b. Dependent Variable: PäättöKA

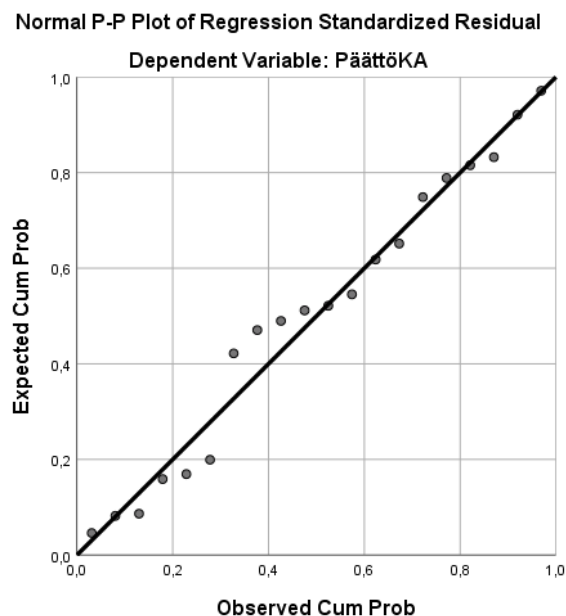
Lineaaristen regressiomallien edeltävyys ehdot on esitetty aiemmin kappaleessa ”5.2 Regressioanalyysi”. Jotta mallin merkitsevyyttä voitaisiin arvioida on nämä ehdot täytettävä. Jäännöskuvioista voidaan päätellä, että suuria lineaarisuuden tai varianssien yhtäsuuruuden rikkomuksia mallissa ei esiinny, sillä pisteet ovat jakautuneet melko satunnaisesti jäännöskuviossa.

Kuvaaja 10. Ensimmäisen mallin jäännösten normaalijakaumakuvio.

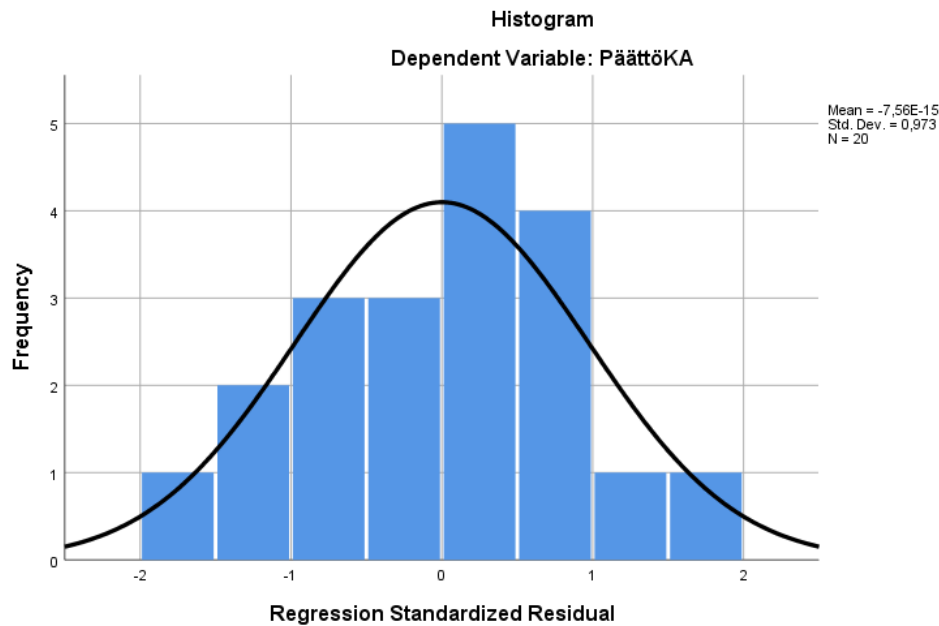


Jäännösten normaalijakautuneisuutta tarkasteltiin normaalijakaumakuvion avulla, kuten Taanila (2020a, 14) suosittelee. Lisäksi havainnot varmistettiin histogrammin avulla. Sekä normaalijakaumakuvioista, että histogrammista voidaan päätellä, että mallin tuottamat jäännöstermit ovat normaalisti jakautuneita: kovin suuria poikkeamia normaalijakaumaa kuvaavalta viivalta ei esiinny ja myös histogrammin avulla voidaan todeta jäännösten normaalijakautuneisuuden olevan mallissa kohtuullisen selkeää.

Kuvaaja 11. Ensimmäisen mallin normaalijakaumakuvio.



Kuvaaja 12. Ensimmäisen mallin jäännösten histogrammi.



Jäännösten riippuvuus saattaa Taanilan (2020a, 15) mukaan muodostua ongelmaksi aikasarjojen kohdalla. Mitkään havainnot eivät kuitenkaan ole tutkimuksessa toisistaan ajallisesti riippuvia. Ehto jäännösten riippumattomuudesta voidaan näin todeta täytetyksi.

Voidaan siis todeta, että mallin edeltävyys ehdot ovat suhteellisen hyvin voimassa. Mallin merkitsevyystestauksia voidaan näin ollen pitää luotettavina. Mallin F-testin p-arvosta (Sig. = ,008) nähdään, että malli on tilastollisesti merkitsevä. Koska mallissa ei ole mukana mitään muuta muuttujaa keskiasteen koulutuksen saaneiden suhteellisen osuuden lisäksi, on selittävän muuttujan korrelaatiokertoimen (Pearson) t-testin tilastollinen merkitsevyys sama kuin koko mallin F-testillä ($r = -,578$, Sig. = ,008), eikä muuttujien selitysosuuksia esittelevää taulukkoa ole näin ollen syytä erikseen tässä tuoda näkyville.

Taulukko 10. Ensimmäisen mallin ANOVA-taulukko.

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,325	1	,325	9,012	,008 ^b
	Residual	,649	18	,036		
	Total	,974	19			

a. Dependent Variable: PäätöKA

b. Predictors: (Constant), työl_kes/työl

Toinen malli

Tutkimuksessa haluttiin muodostaa taustateoriaan tukeutuen myös toinen malli, joka huomioisi alueen tulotason. Tässä toisessa mallissa päädyttiin ottamaan mukaan lähialue-estimaattien koulutustasomuuttujista parhaiten yläkoulujen päättökeskiarvoja selittävä, sekä estimaattien tulotasoja kuvaavista muuttujista parhaiten päättökeskiarvoja selittävä muuttuja. Malliin valikoituivat näin lukuisten testausten kautta keskiasteenkoulutuksen käyneiden suhteellinen osuus, sekä rahatulot per henkilö. Malli muodostettiin käyttämällä SPSS:n Enter-algoritmia, joka ottaa mukaan malliin kaikki tutkijan valitsevat selittävät muuttujat, paitsi jos niissä ilmenee voimakasta multikolinearisuutta.

Edellä kuvatulla tavalla muodostetussa mallissa korjatuksi selityskertoimeksi (Adjusted R²) saatiin 0,317. Tämä tarkoittaa, että mallin avulla yläkoulujen päättökeskiarvosta pystyttäisiin selittämään jopa 31,7 %. Malli saa F-testissä p-arvoksi luvun 0,015, mikä tarkoittaa, että se olisi tilastollisesti merkitsevä, mikäli merkitsevyytestauksiin voidaan luottaa. Mallin merkitsevyydestä on kuitenkin vaikea pitää pätevänä, sillä normaalijakaumakuvion ja histogrammin avulla mallin tuottamat jäännökset eivät noudata kovin hyvin normaalijakaumaa. Sen sijaan mallin lineaarisuutta tai jäännösten varianssia ei liene syytä epäillä mallin jäännöskuvion perusteella. Yksittäisistä muuttujista keskiasteen koulutuksen saaneiden osuutta kuvaavan muuttujan yhteys päättökeskiarvoon oli mallissa kohtuullinen ja tilastollisesti merkitsevä ($r = -,470$, Sig. = ,038). Sen sijaan rahatuloja per henkilö kuvaavan muuttujan yhteys päättökeskiarvoon oli heikko, eikä yhteys saavuttanut korrelaatiokertoimen t-testissä tilastollista merkitsevyyttä ($r = ,258$, Sig. = ,233). Näiden tietojen perusteella voidaan todeta, että mallin tilastollinen merkitsevyys on kyseenalainen, eikä mallin antamaa selityskerrointa ja -astetta voida pitää kovinkaan luotettavana.

Taulukko 11. Toisen mallin yhteenveto.

Model Summary ^b				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,623 ^a	,389	,317	,18716

a. Predictors: (Constant), rahatulo/henk, tyol_kes/tyol

b. Dependent Variable: PäättöKA

Taulukko 12. Toisen mallin ANOVA-taulukko.

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,379	2	,189	5,404	,015 ^b
	Residual	,596	17	,035		
	Total	,974	19			

a. Dependent Variable: PäättöKA

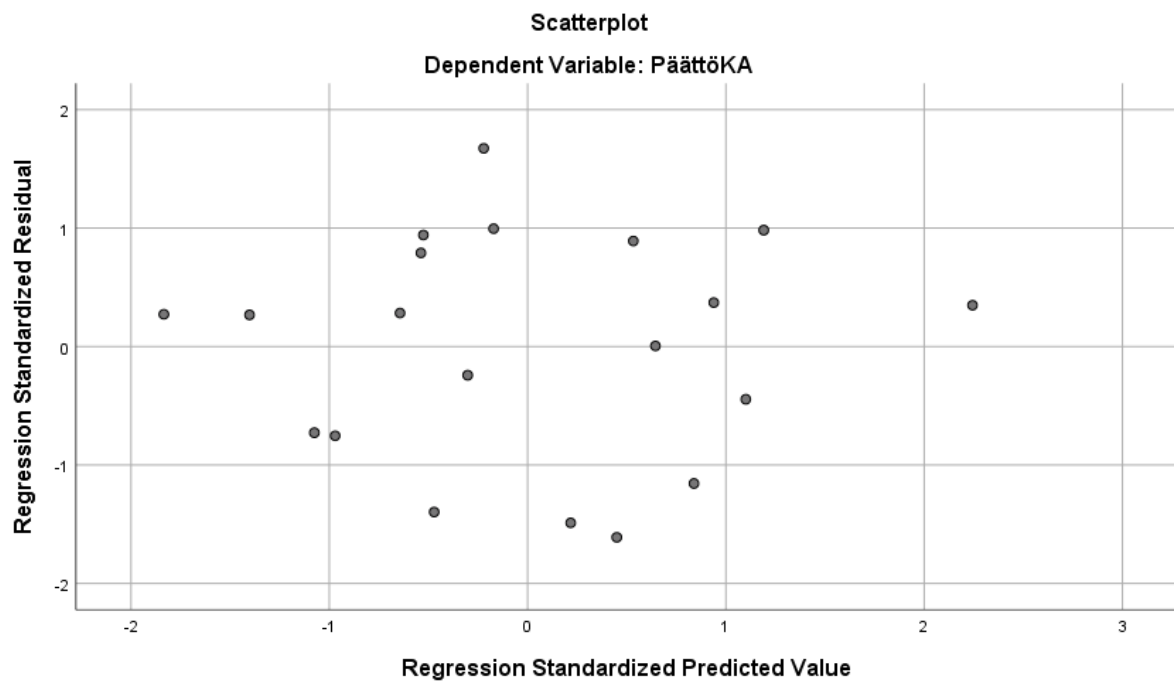
b. Predictors: (Constant), rahatulo/henk, tyol_kes/tyol

Taulukko 13. Toisen mallin muuttujien erittely.

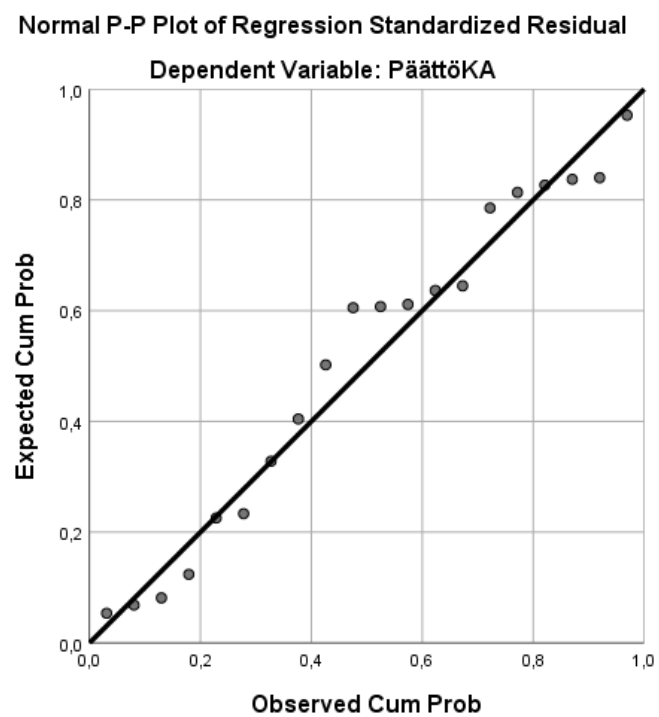
Coefficients ^a										
		Unstandardized Coefficients		Standardized Coefficients			95,0 % Confidence Interval for B		Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	8,206	,728		11,277	,000	6,671	9,741		
	tyol_kes/tyol	-,016	,007	-,470	-2,256	,038	-,030	-,001	,827	1,209
	rahatulo/henk	3,589E-5	,000	,258	1,237	,233	,000	,000	,827	1,209

a. Dependent Variable: PäättöKA

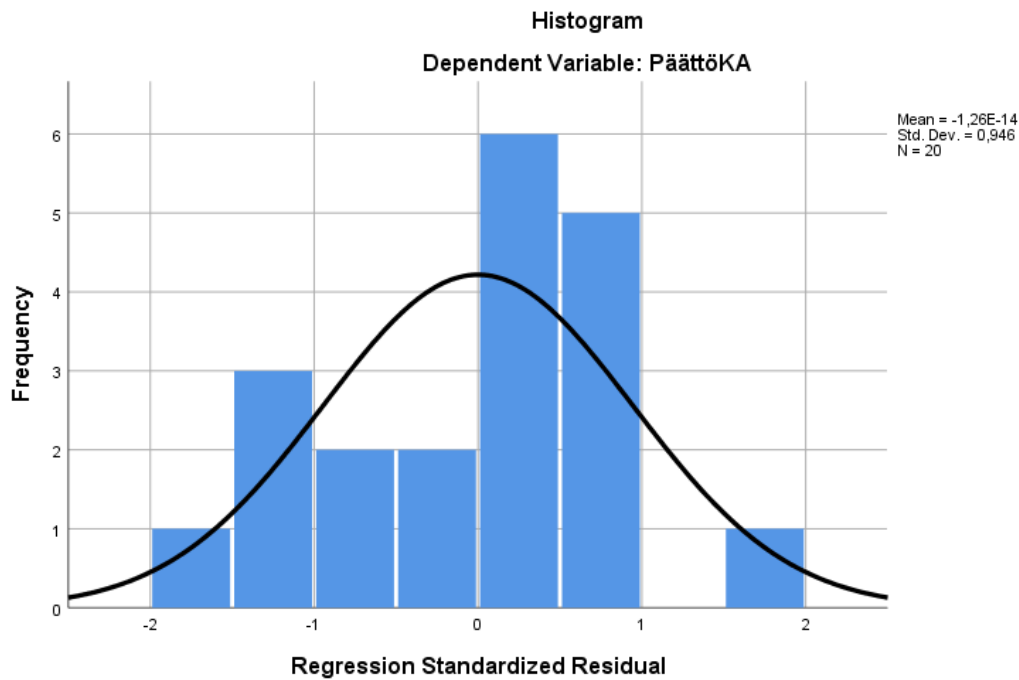
Kuvaaja 13. Toisen mallin jäännösten hajontakuvio.



Kuvaaja 14. Toisen mallin jäännösten normaalijakaumakuvio.



Kuvaaja 15. Toisen mallin jäännösten histogrammi.



4.2 Yläkoulujen lähialue-estimaatit oppilasalueiden kanssa

Lähialue-estimaattien analyysi toteutettiin tutkimuksessa kahdella tavalla, joista toisessa hyödynnettiin tietoa siitä, missä kaupungin määrittelemät oppilasalueiden rajat kulkevat. Todellisessa tilanteessa oppilaita siirtyy myös rajojen yli, mutta analyysissä oletettiin tilanne, jossa koulua käydään vain oman oppilasalueen sisällä. Lähialue-estimaatit rajattiin oppilasalueiden rajoihin ja tällä menetelmällä saadut tulokset esitetään tässä luvussa. Luvun ensimmäinen alaluku esittää tuloksissa havaittuja työssäkäyvän työväestön koulutustason yhteyksiä oppimistuloksiin, eli yläkoulujen päättökeskiarvoihin. Toisessa alaluvussa esitetään puolestaan tulokset, jotka käsittelevät lähialue-estimaattien laskennallisten tulosten yhteyksiä yläkoulujen päättökeskiarvoihin. Kolmannessa alaluvussa esitetään erilaisten regressiomallien tulokset tilanteessa, jossa lähialue-estimaatteja rajaavat sivistys- ja kulttuurialueet.

Koska Pearsonin korrelaation toteuttamisen ehtojen ei voitu katsoa olevan voimassa myöskään sivistys- ja kulttuurialueiden rajaamien lähialue-estimaattien kanssa, päädyttiin analyysit toteuttamaan Spearmanin järjestyskorrelaatiokertoimen avulla.

4.2.1 Työväestön koulutus ja oppimistulokset

Spearmanin järjestyskorrelaatiokerrointa käyttävän korrelaatioanalyysin avulla selvisi, että korrelaation voimakkuus yläkoulujen päättökeskiarvojen ja pelkän perusasteen koulutuksen saaneiden osuuden työssäkäyvistä työvoimasta välillä oli korkeintaan kohtuullinen ($r_s = -,394$) ja negatiivinen, eikä myöskään tilastollisesti merkitsevä. Pelkän peruskoulun suorittaneen työvoiman osuuden kasvaessa päättökeskiarvo yläkoululla siis laskee ja toisinpäin. Kaikista voimakkainta korrelaatiota päättökeskiarvon kanssa kolmesta koulutustasosta kuvaavasta muuttujasta osoitti keskiasteen koulutuksen saaneen työssäkäyvän työvoiman suhteellinen osuus. Tässäkin tapauksessa yhteys oli negatiivinen ($r_s = -,530$), ollen myös tilastollisesti merkitsevä (Sig. = ,011). Korkeakoulutetun työvoiman ollessa kyseessä korrelaatio jäi hieman heikommaksi ($r_s = ,505$), mutta oli kuitenkin korkeaa ja tilastollisesti merkitsevää (Sig. = ,017).

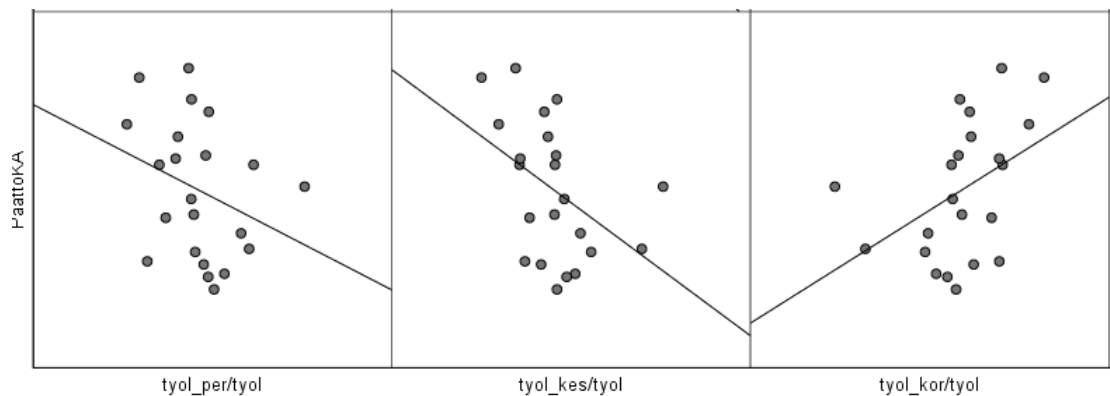
Taulukko 14. Päättökeskiarvon ja koulutusmuuttujien väliset korrelaatiot.

Correlations			PäättöKA	tyol_per/tyol	tyol_kes/tyol	tyol_kor/tyol
Spearman's rho	PäättöKA	Correlation Coefficient	1,000	-,394	-,530*	,505*
		Sig. (2-tailed)	.	,069	,011	,017
		N	22	22	22	22

*. Correlation is significant at the 0.05 level (2-tailed).

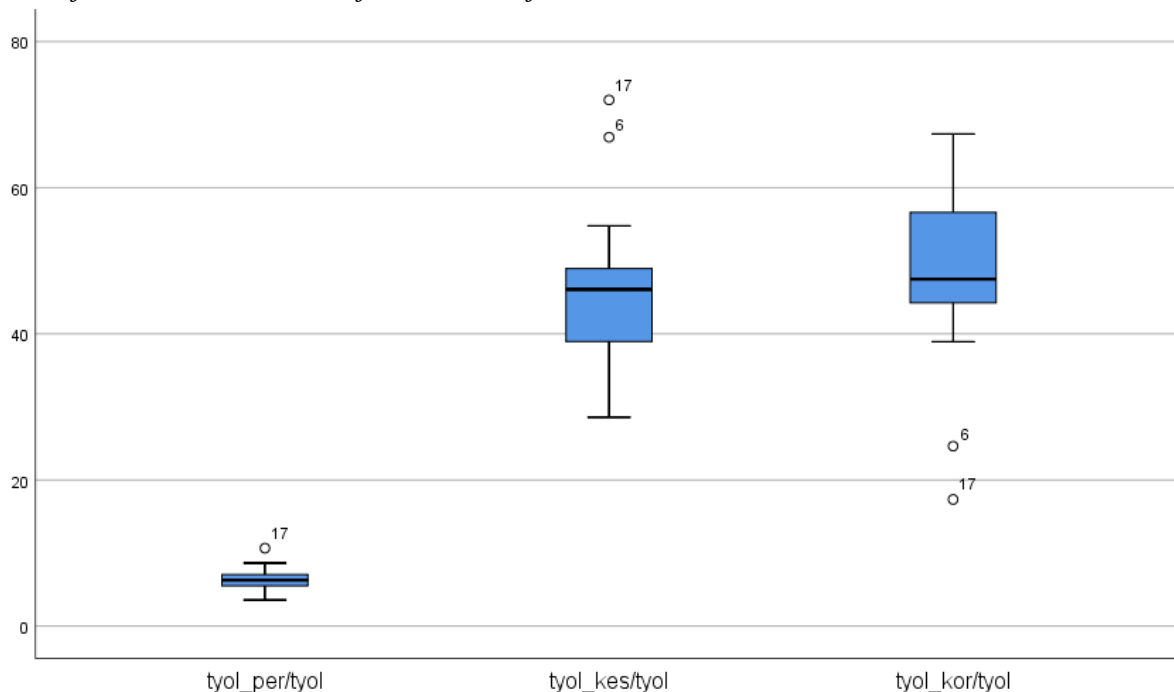
Korrelaation hajontakuvioiden ja niihin lisättyjen regressiosuorien avulla pystyttiin toteamaan samat asiat kuin korrelaatioanalyysin keinoin. Lisäksi ilmeni syytä epäillä aineistossa olevan mukana poikkeavia havaintoja, joita päätettiin tutkia laatikko-janakuvion avulla.

Kuvaaja 16. Päättökeskiarvon ja koulutusmuuttujien hajontakuviot.



Laatikko-janakuvion avulla varmistettiin aineistossa olevan poikkeavia havaintoja koulutustasoa kuvaavissa muuttujissa. Aineistoa tutkittaessa todettiin, että kyseiset havainnot ovat luotettavuudeltaan niin heikkoja, että ne on järkevää poistaa aineistosta ja suorittaa korrelaatioanalyysi uudestaan.

Kuvaaja 17. Koulutusmuuttujien laatikko-janakuviot.



Työväestön koulutus ja päättökeskiarvot ilman poikkeavia havaintoja

Kun korrelaatioanalyysi suoritettiin uudestaan siten, että poikkeavia arvoja sisältävät havaintopisteet oli poistettu aineistosta, havaittiin päättökeskiarvon ja selittävien muuttujien välisten korrelaatioiden kasvaneen. Voimakkain (negatiivinen) yhteys löytyi edelleen

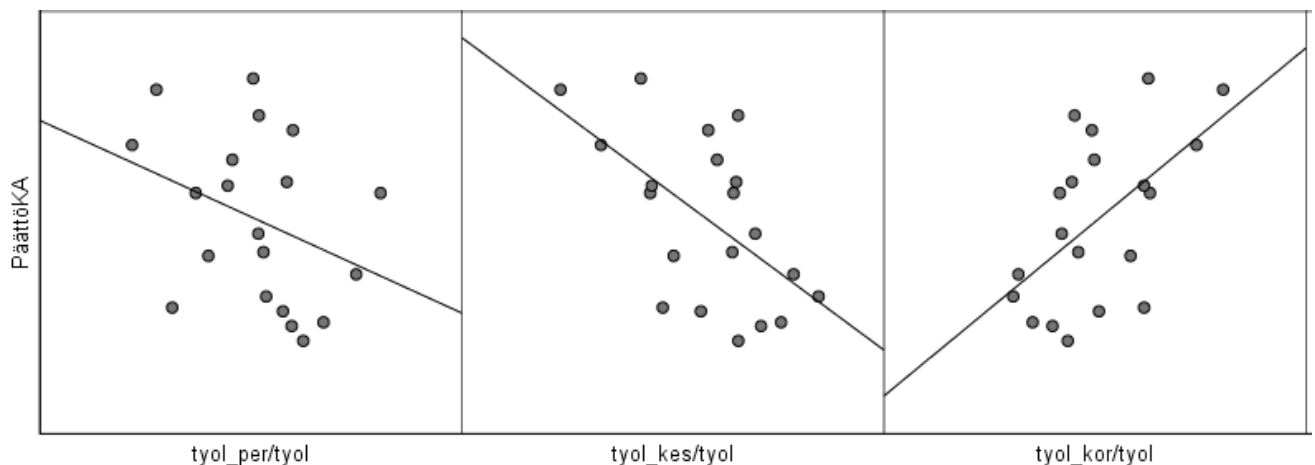
keskiasteen koulutuksen saaneiden osuutta kuvaavan muuttujan ja päättökeskiarvon väliltä, ollen myös tilastollisesti merkitsevä 0.01 tasolla ($r_s = -,564$, Sig. = ,010). Positiivinen tilastollisesti merkitsevä yhteys löytyi päättökeskiarvon ja korkeakoulutuksen käyneiden osuuden väliltä ($r_s = ,538$, Sig. = ,014). Negatiivinen melko vahva yhteys löytyi myös pelkän peruskoulun käyneiden osuuden ja päättökeskiarvon väliltä ($r_s = -,424$), mutta tämä ei kuitenkaan ollut tilastollisesti merkitsevää (Sig. = ,062).

Taulukko 15. Päättökeskiarvon ja koulutusmuuttujien väliset korrelaatiot.

			tyol_per/tyol	tyol_kes/tyol	tyol_kor/tyol
Spearman's rho	PäättöKA	Correlation Coefficient	-,424	-,564**	,538*
		Sig. (2-tailed)	,062	,010	,014
		N	20	20	20

Myös hajontakuvioiden avulla voitiin silmämääräisesti todeta sama asia, minkä korrelaatioanalyysit osoittivat: yhteyden voimakkuus päättökeskiarvojen ja koulutustasoa kuvaavien muuttujien välillä voimistui, kun poikkeavia havaintoja sisältävät havaintopisteet poistettiin tarkastelusta kokonaan.

Kuvaaja 18. Päättökeskiarvon ja koulutusmuuttujien väliset hajontakuviot.



4.2.2 Lähialueen tulot ja oppimistulokset

Sivistys- ja kulttuurialueiden rajaamien lähialue-estimaattien tulojen ja päättökeskiarvojen korrelaatioita tutkittaessa päädyttiin jälleen käyttämään Spearmanin järjestyskorrelaatiokerrointa, sillä muuttujien ei voitu todeta olevan riittävän normaalisti

jakauneita Pearsonin korrelaatiokerrointa varten. Lisäksi haluttiin vertailla tuloksia SIKU-alueiden rajaamien ja rajaamattomien lähialue-estimaattien välillä, minkä vuoksi Spearmanin järjestyskorrelaatiokertoimen käyttö oli järkevää.

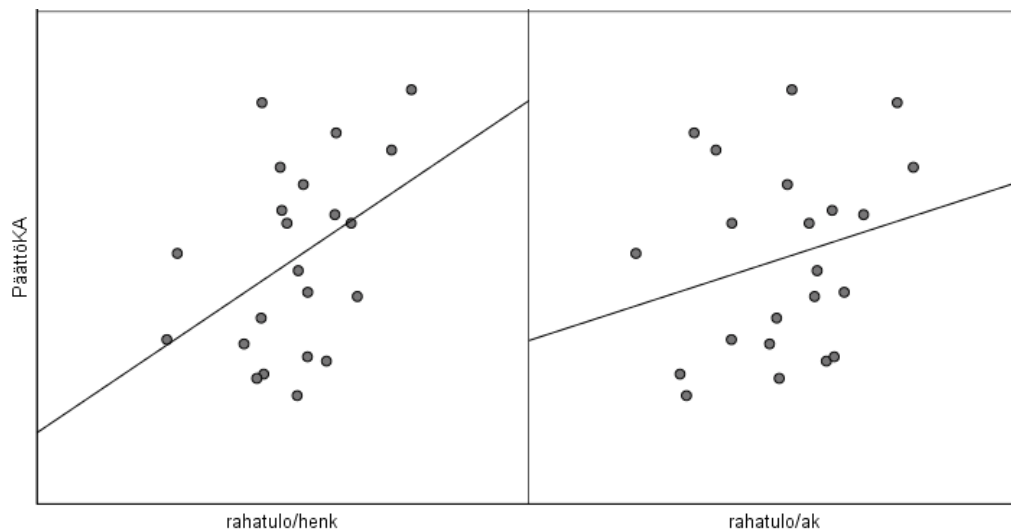
Korrelaatioanalyysi osoitti, että kovin merkittäviä yhteyksiä ei lähialue-estimaattien tuloja kuvaavien muuttujien ja yläkoulujen päättökeskiarvojen väliltä löydy. Rahatuloja per henkilö kuvaavalla muuttujalla yhteys ($r_s = ,384$, Sig. = $,078$) oli hieman vahvempi kuin asuntokuntaakohtaisia tuloja kuvaavan muuttujan kohdalla ($r_s = ,264$, Sig. = $,238$). Kummankaan muuttujan yhteys yläkoulujen päättökeskiarvoihin ei ollut tilastollisesti merkitsevää.

Taulukko 16. Päättökeskiarvon ja tulomuuttujien väliset korrelaatiot.

			rahatulo/henk	rahatulo/ak
Spearman's rho	PäättöKA	Correlation Coefficient	,384	,264
		Sig. (2-tailed)	,078	,236
		N	22	22

Hajontakuvioiden avulla nähtiin, että havainnot todellakin ovat melko satunnaisesti hajonneita, eikä kovin suurta yhteyttä muuttujien väliltä voi löytää silmämääräisesti tarkasteltuna. Hajontakuvioiden avulla ei ollut myöskään syytä epäillä, että muuttujat sisältäisivät poikkeavia havaintoja. Myös laatikko-janakuvioiden avulla todennettiin, että muuttujat eivät sisältäneet poikkeavia havaintoja.

Kuvaaja 19. Päätökeskiarvon ja tulomuuttujien väliset hajontakuviot.



Työväestön tulojen ja päätökeskiarvojen yhteys ilman epäluotettavia havaintopisteitä

Tutkimusaineiston tuntemisen perusteella oli järkevää suorittaa korrelaatioanalyysi myös siten, että muutama havaintopiste oli poistettu aineistosta, vaikka ne eivät näyttäytyneetkään poikkeavina havaintoina laatikko-jana- tai hajontakuvioiden. Tämä johtuu siitä, että kyseisten havaintojen luotettavuus on kyseenalaistettava niiden sisältäessä vain erittäin pienen määrän oman lähialue-estimaattinsa sisällä todellisuudessa olevista tiedoista. Kyseisten havaintojen tapauksessa määrät olivat jopa alle 5 % alueen tilastotietojen kokonaismääristä. Näin ollen kyseiset kaksi havaintoa päädyttiin poistamaan aineistosta ja suorittamaan korrelaatioanalyysi uudelleen tällä tavalla.

Analyysin perusteella merkittävää muutosta korrelaatiokertoimissa ei tällä menetelmällä tapahtunut verrattuna aikaisempaan. Rahatulojen per henkilö yhteys päätökeskiarvoon pysyi täsmälleen samana ($r_s = ,384$, Sig. = ,094) verrattuna edelliseen analyysiin, mutta sai hieman heikomman tuloksen merkitsevyystestissä. Rahatulojen per asutuskunta kohdalla merkitsevyystestin tulos (Sig. = ,243) kasvoi myös, mikä tarkoittaa, että merkitsevyys päättyi kauemmas tilastollisen merkitsevyyden rajana yleisesti pidetystä p-arvosta 0,05. Merkitsevyyden heikkenemisen lisäksi asutuskuntaakohtaisia tuloja kuvaavan muuttujan yhteys päätökeskiarvoon kasvoi aavistuksen ($r_s = ,274$). Havaintojen määrä ($N = 20$) oli tässä versiossa testiä pienempi kuin aiemmassa, kaikki havainnot käsittävässä muuttujassa, mikä todennäköisesti selittää merkitsevyyden heikentymisen. Erot edellisen testin kanssa olivat niin pieniä, että tässä yhteydessä ei esitetä jälkimmäiseen testiin liittyviä hajontakuviota.

Taulukko 17. Päätökeskiarvon ja tulomuuttujien väliset korrelaatiot.

Correlations

			PäätöKA	rahatulo/henk	rahatulo/ak
Spearman's rho	PäätöKA	Correlation Coefficient	1,000	,384	,274
		Sig. (2-tailed)	.	,094	,243
		N	20	20	20

4.2.3 Regressioanalyysien tulokset

Kuten aiemmassa luvussa SIKU-alueilla rajaamattomien lähialue-estimaattien tapauksessa, myös SIKU-alueilla rajattujen estimaattien kanssa päädyttiin muodostamaan kaksi erilaista mallia. Ensimmäinen malli muodostettiin SPSS:n Stepwise-algoritmin avulla sellaiseksi, että se sisältää parhaiten selitettävän muuttujan vaihtelua selittävät muuttujat. Toiseen malliin otettiin puolestaan tutkimuksen taustateoriaan liittyvistä syistä mukaan päätökeskiarvon vaihtelua parhaiten selittävä rahatuloja kuvaava, sekä alueen työväestön koulutustasoa kuvaava muuttuja.

Malleja muodostettaessa päädyttiin jättämään pois samat kaksi havaintoa, jotka jätettiin pois jo aiemmissa regressioanalyysissä. Perusteluna toimii aineiston tunteminen eli se, että näiden havaintojen kohdalla ei voida pitää koulutustasoa tai rahatuloja kuvaavia muuttujia luotettavina niiden sisältämän pienen tietomäärän vuoksi. Lisäksi kyseiset havainnot ilmenivät poikkeavina havaintoina korrelaatiotestauksien yhteydessä koulutusmuuttujien laatikko-janakuvioissa.

Ensimmäinen malli

Ensimmäisen SIKU-alueiden sisällä muodostettujen lähialue-estimaattien perusteella muodostetun mallin tulokset eivät poikenneet paljonkaan ilman SIKU-alueita muodostetun parhaan mallin tuloksista. Muuttujia jäi Stepwise-algoritmin tuottamaan malliin vain yksi: keskiasteen koulutuksen saaneen työssäkäyvän työväestön suhteellinen osuus lähialue-estimaatin sisällä asuvasta työssäkäyvästä työvoimasta. Havaintoja mallissa oli mukana 20 kappaletta koska muutamia havaintoja poistettiin.

Mallin tuottama korjattu selityskertoimen ($\text{Adjusted } R^2 = ,298$) mukaan parhaan mahdollisen SIKU-alueiden sisäisen mallin selitysasteeksi muodostuu 29,8 % päätökeskiarvon vaihtelusta.

korrelaatiokertoimen arvo mallissa oli 0,579 ja se on samalla myös ainoan selittävän muuttujan saama arvo. Samoin mallin luotettavuustestauksen tuottama arvo 0,007, joka on tilastollisesti merkitsevä, on yhtäläinen muuttujan t-testin arvon kanssa. Mallin edeltävyysehtoja tarkastellessa huomataan jäännöskuvion avulla jäännösten varianssin olevan melko hyvin samassa suuruusluokassa. Samoin voidaan havaita, että mallin muuttujien välistä lineaarisuutta ei ole syytä epäillä. Jäännösten voi todeta myös noudattavan hyvin normaalijakaumaa sekä normaalijakaumakuvion, että histogrammin kautta tarkasteltuna. Lisäksi tiedetään, että jäännökset ovat toisistaan riippumattomia, joten voidaan todeta mallin edeltävyysehtojen olevan hyvin voimassa ja merkitsevyystestauksen olevan pätevä.

Taulukko 18. Ensimmäisen mallin yhteenveto.

Model Summary ^b				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,579 ^a	,335	,298	,18968

a. Predictors: (Constant), tyol_kes/tyol

b. Dependent Variable: PäättöKA

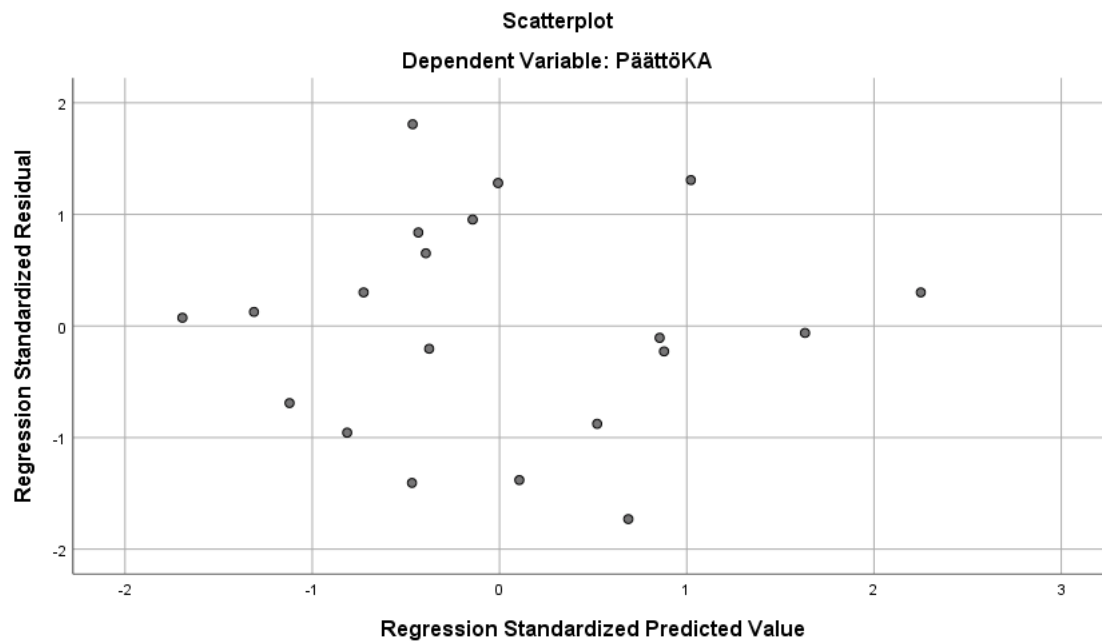
Taulukko 19. Ensimmäisen mallin ANOVA-taulukko.

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,326	1	,326	9,074	,007 ^b
	Residual	,648	18	,036		
	Total	,974	19			

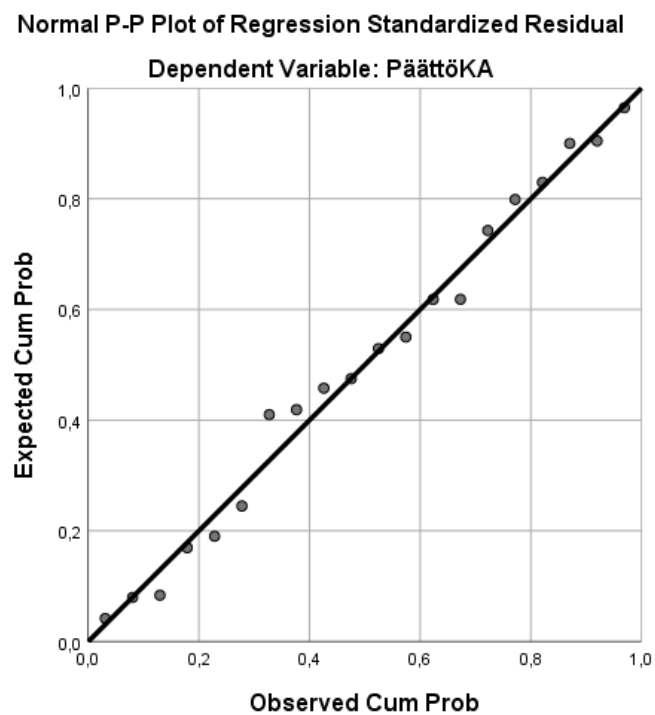
a. Dependent Variable: PäättöKA

b. Predictors: (Constant), tyol_kes/tyol

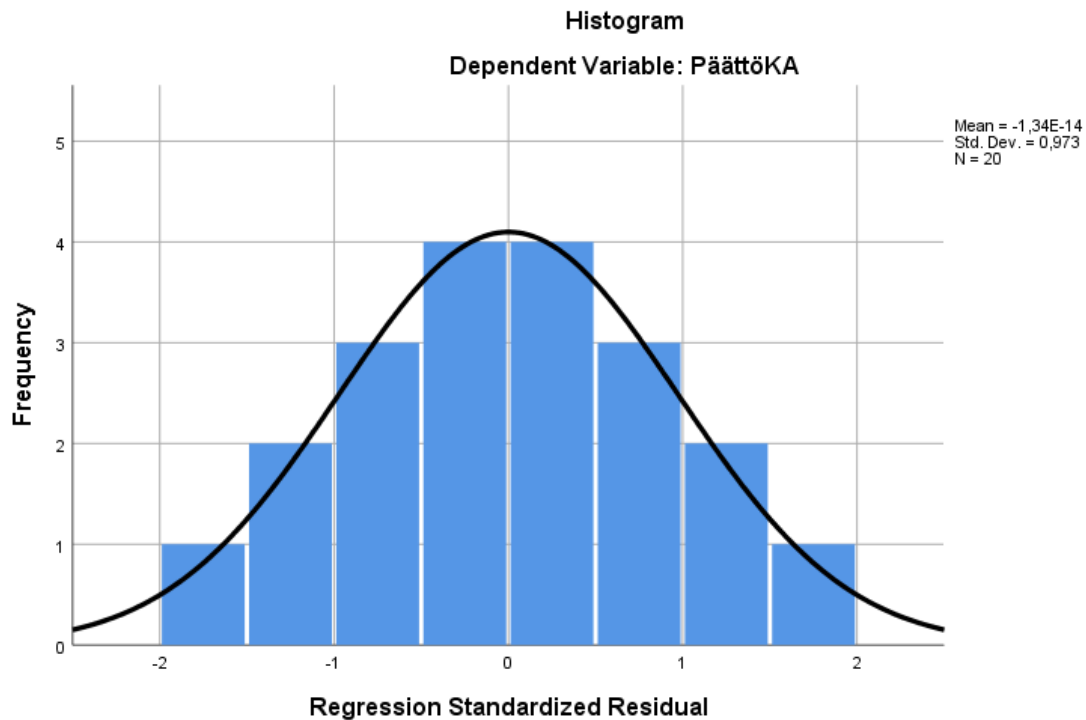
Kuvaaja 20. Ensimmäisen mallin jäännösten hajontakuvio.



Kuvaaja 21. Ensimmäisen mallin jäännösten normaalijakaumakuvio.



Kuvaaja 22. Ensimmäisen mallin jäännösten histogrammi.



Toinen malli

Toisessa mallissa haluttiin muodostaa taustateoriaan pohjautuen malli, joka huomioi myös parhaiten päätöskeskivärtien vaihtelua selittävän tulomuuttujan, vaikka tiedossa oli, että tämä heikentää mallin merkitsevyyttä. Malliin valikoituivat muuttujiksi erilaisten yhdistelmien testaamisen kautta keskiasteen koulutuksen käyneiden osuus alueen työssäkäyvistä työvoimasta sekä laskennallisia rahatuloja per henkilö kuvaava muuttuja. Havaintoja (N) mallissa oli mukana 20 kappaletta.

Mallin selitysaste nousi 32,9 prosenttiin (Adjusted $R^2 = ,329$) sen F-testin p-arvon ollessa 0,013 ja ollen näin tilastollisesti merkitsevä. Merkitsevyytestauksen pätevyyttä pohdittaessa huomataan, että mallin jäännöskuvio ei anna aihetta epäillä selitettävän ja selittävien muuttujien välisiä lineaarisuuksia, eikä hyvin epäsäännöllisesti muodostunut kuvio anna epäillä myöskään jäännösten varianssien yhtäsuuruusehdon rikkomista. Sen sijaan jäännösten normaalijakaumakuvio antaa aihetta epäillä jäännösten normaalijakautuneisuutta, sillä osa pisteistä poikkeaa viivalta vahvasti. Histogrammin avulla voidaankin todeta, että jäännösten normaalijakautuneisuus ei mallissa täyty kovin hyvin, sillä histogrammi on vahvasti oikealle

nousevasti vino. Näin ollen todetaan, että mallin edeltävyys ehdot eivät ole voimassa, eikä merkitsevyystestaukselle voi antaa kovin suurta painoarvoa mallia arvioitaessa.

Yksittäisistä muuttujista vahvinta yhteyttä päättökeskiarvojen kanssa ilmaisi keskiasteen koulutuksen käyneiden osuus, korrelaation suunta oli negatiivinen, melko vahva ja tilastollisesti merkitsevä ($\beta = -,475$, Sig = ,032). Rahatulomuuttujan kohdalla yhteys oli heikkoa, positiivista ja ei tilastollisesti merkitsevää ($\beta = ,274$, Sig = ,194).

Taulukko 20. Toisen mallin yhteenveto

Model Summary ^b				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,632 ^a	,400	,329	,18546

a. Predictors: (Constant), tyol_kes/tyol, rahatulo/henk

b. Dependent Variable: PäättöKA

Taulukko 21. Toisen mallin ANOVA-tilaus.

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,389	2	,195	5,661	,013 ^b
	Residual	,585	17	,034		
	Total	,974	19			

a. Dependent Variable: PäättöKA

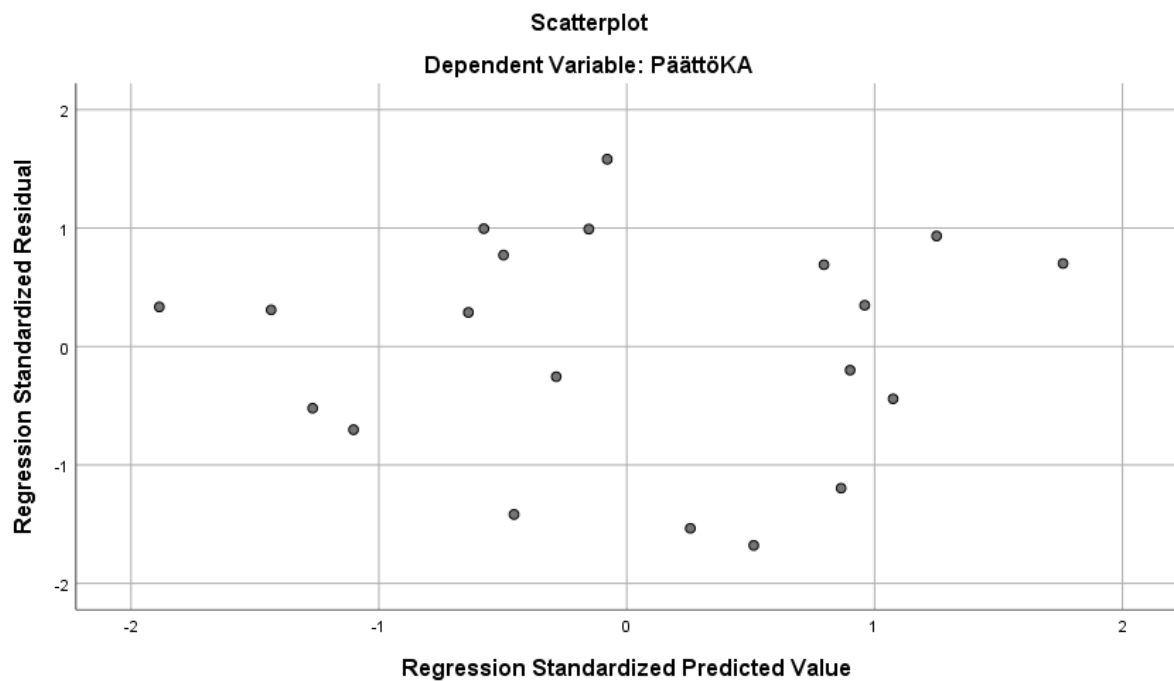
b. Predictors: (Constant), tyol_kes/tyol, rahatulo/henk

Taulukko 22. Toisen mallin muuttujien erittely.

Coefficients ^a										
		Unstandardized Coefficients		Standardized Coefficients			95,0% Confidence Interval for B		Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	8,170	,721		11,335	,000	6,649	9,690		
	rahatulo/henk	3,938E-5	,000	,274	1,352	,194	,000	,000	,857	1,167
	tyol_kes/tyol	-,016	,007	-,475	-2,342	,032	-,031	-,002	,857	1,167

a. Dependent Variable: PäättöKA

Kuvaaja 23. Toisen mallin jäännösten hajontakuviio.



5 Tulosten koonti ja johtopäätökset

Seuraavaksi esitellään kootusti taulukoissa tutkimuksen tulosten ja johtopäätösten muodostamisen kannalta keskeisimmät tulokset korrelaatio- ja regressioanalyyseistä. Ensimmäisenä esitetään tulokset, kun lähialue-estimaattien muodostumista ei ole rajattu sivistys- ja kulttuurialueiden rajoja hyödyntämällä (Taulukko 23./”OULU”). Toisessa taulukossa puolestaan esitetään tulokset tilanteessa, jossa rajat on otettu huomioon lähialue-estimaatteja luotaessa (Taulukko 24./”SIKU”). Keskeisimmät tulokset esitellään tekstiosiossa taulukoiden jälkeen. Tulosten koonnin jälkeen niitä tarkastellaan suhteessa tutkimuksen teoriaan ja samalla tehdään tuloksista johtopäätökset.

Taulukko 23. Ilman SIKU-alueita muodostettujen lähialueiden tulokset.

OULU						
Korrelaatioanalyysit						
Päättökeskiarvoa selittävä muuttuja	Kaikki havainnot mukana			Poikkeavat havainnot poistettu		
	Spearman's rho	Sig.	Havaintojoukko N	Spearman's rho	Sig.	Havaintojoukko N
työl_per	-,388	,074	22	-,408	,074	20
työl_kes	-,515*	,014	22	-,547*	,013	20
työl_kor	,488*	,022	22	,517*	,020	20
rahatulo/henk	,377	,084	22	,272	,260	19
rahatulo/ak	,245	,272	22	Ei poikkeavia havaintoja	-	-
*tilastollisesti merkitsevä yhteys						
Regressioanalyysit						
Regressiomalli (muuttujat)	Korjattu selityskerroin		Merkitsevyys			
Malli 1 (työl_kes)	,297		,008			
Malli 2* (työl_kes + rahatulo_henk)	,317		,015*			
*mallin edeltävyys ehdot eivät täyty eli merkitsevyystestausta ei voida pitää luotettavana						

Taulukko 24. SIKU-alueita hyödyntäen muodostettujen lähialueiden tulokset.

SIKU						
Korrelaatioanalyysit						
Päättökeskiarvoa selittävä muuttuja	Kaikki havainnot mukana			Poikkeavat havainnot poistettu		
	Spearman	Sig.	Havaintojoukko N	Spearman	Sig.	Havaintojoukko N
työl_per/työl	-,394	,069	22	-,424	,062	20
työl_kes/työl	-,530*	,011	22	-,564*	,010	20
työl_kor/työl	,505*	,017	22	,538*	,014	20
rahatulo/henk	,384	,078	22	,384	,094	20
rahatulo/ak	,264	,236	22	,274	,243	20
*tilastollisesti merkitsevä yhteys						
Regressioanalyysit						
Regressiomalli		Korjattu selitysaste		Merkitsevyys		
Malli 1 (työl_kes)		,298		,007		
Malli 2* (työl_kes + rahatulo_henk)		,329		,013*		
*mallin edeltävyys ehdot eivät täyty eli merkitsevyystestausta ei voida pitää luotettavana						

Tutkimuksessa toteutetut korrelaatioanalyysit osoittivat, että yläkoulujen päättökeskiarvoihin voimakkaimmin yhteydessä olivat lähialue-estimaatin sosioekonomisia tekijöitä kuvaavista muuttujista alueen työssäkäyvän väestön koulutustasoa kuvaavat muuttujat. Näistä muuttujista vahvasti päättökeskiarvoon yhteydessä olivat muuttujat, jotka kuvasivat korkeintaan keskiasteen koulutuksen saaneen työväestön osuutta kaikista alueella asuvista työssäkävijöistä *työl_kes/työl*, sekä korkeakoulutuksen saaneen työväestön osuutta kuvaava muuttuja *työl_kor/työl*. Tosin myös pelkän perusasteen koulutuksen saaneen työvoiman osuudella *työl_per/työl* havaittiin yhteys päättökeskiarvoon, mutta missään tarkasteluissa kyseinen muuttuja ei kuitenkaan osoittautunut tilastollisesti merkitseväksi (Sig. < ,050). Tarkastellessa korrelaatioita tarkemmin kahden vahvimmin päättökeskiarvoon yhteydessä olevan muuttujan osalta huomataan, että kyseiset muuttujat saavuttivat tilastollisen merkittävyyden kaikissa eri tavoin toteutetuissa korrelaatioanalyysissä. Voidaan siis melko luotettavasti todeta, että yhteys näiden muuttujien ja yläkoulun päättökeskiarvon välillä on todella olemassa.

Vahvimmin päättökeskiarvon kanssa korreloivien muuttujien *työl_kes/työl* ja *työl_kor/työl* kohdalla havaittiin myös, että ne korreloivat myös keskenään hyvin voimakkaasti. Kun yhdistetään tähän tieto siitä, että muuttujat saivat korrelaatioanalyysissä eri suuntaisen yhteyden päättökeskiarvon kanssa toisiinsa nähden, voidaan todeta, että muuttujat kuvaavat oikeastaan samaa asiaa. Jostain syystä kuitenkin päättökeskiarvon yhteys oli voimakkaampi nimenomaan keskiasteen koulutusta kuvaavan muuttujan kanssa. Yhteys oli negatiivinen ja pysyi saman suuntaisena jokaisessa eri tavalla toteutetussa korrelaatioanalyysissä. Käytännössä tämä tarkoittaa sitä, että korkeintaan keskiasteen koulutuksen saaneen työväestön osuuden noustessa alueen työssäkäyvien asukkaiden joukossa, voidaan yläkoulujen päättökeskiarvojen olettaa laskevan. Vastaavasti korkeakoulutettujen osuuden noustessa työssäkäyvistä työväestöstä alueen yläkoulun päättökeskiarvojen voi todeta nousevan. Yleistäen nämä tulokset pysyivät samoina, riippumatta siitä rajattiinko lähialue-estimaattien muodostumista vai ei. Myöskään poikkeavien havaintojen poistamisella ei ollut kovin suurta merkitystä, mutta poikkeavien havaintojen poistaminen nosti korrelaatiokertoimien saamia merkitsevyyksiä aavistuksen kummassakin tapauksessa, samoin nosti korrelaation voimakkuutta hieman.

Rahatuloja kuvaavien muuttujien osalta *rahatulo/ak* – joka kuvaa asutokunnan keskimääräisiä käytössä olevia rahatuloja koulun lähialueella – osalta yhteys yläkoulun päättökeskiarvoon oli heikko, sillä arvot eri skenaarioista vaihtelivat välillä $r_s = 0,245\text{--}0,274$, eikä myöskään tilastollista merkitsevyyttä muuttujien väliltä pystytty osoittamaan missään eri tavoin toteutetussa korrelaatioanalyysissä. Edes keskimääräisiä rahatuloja per henkilö kuvaava muuttuja *rahatulo/henk* ei osoittautunut missään analyysissä tilastollisesti merkittäväksi. Korrelaatio henkilöä kohti laskettujen keskimääräisten rahatulojen ja päättökeskiarvojen välillä oli kuitenkin hieman vahvempi kuin asutokuntakohtaisia tuloja kuvaavan muuttujan kohdalla korrelaatiokertoimen arvon vaihdellessa välillä $0,272\text{--}0,384$.

Korrelaatioanalyysien perusteella voidaan todeta, että Oulussa esiintyy koulujen oppimistulosten eriytymistä, joka on yhteydessä lähialueen sosioekonomisiin tekijöihin. Voimakkainta yhteyttä osoittaa koulun lähialueen työssäkäyvän väestön koulutustaso: Mitä korkeampi on väestön koulutustaso koulun lähialueella, sitä parempia oppimistuloksia koululta oli odotettavissa. Tulos vastaa hyvin aiempien tutkimusten tuloksia (esim. Bernelius, 2005; 2010; 2013a). Berneliuksen (2005, 47; 2013a, 108, 139–140 ja 178) tutkimuksissaan pääkaupunkiseudun alueilta löytämät korrelaatiot korkeakoulutettujen aikuisten alueellisen osuuden ja koulujen oppimistulosten väliltä sekä suomenkielisten ylä- ($r = .790, p < .001$),

että alakoulujen ($r = .740, p < .001$) osalta ovat hieman vahvempia kuin tässä tutkimuksessa saadut.

Berneliuksen (2005, 47; 2013a, 108) tutkimuksissa heikosti koulutettujen – eli vain peruskoulun käyneiden – aikuisten osuudella oli lähes yhtä vahva, mutta negatiivinen yhteys oppimistuloksiin kuin korkeakoulutettujen osuudella. Tässä tutkimuksessa puolestaan yhteys peruskoulun käyneen työllisen työvoiman osuuden ja oppimistulosten välillä jäi huomattavasti heikommaksi. Syy poikkeavuuteen saattaa löytyä esimerkiksi kaupunkirakenteesta, tutkimuksen toteutustavasta tai aineistoista.

Myös alueen asukkaiden tuloilla näyttäisi tutkimuksen tulosten perusteella olevan heikko yhteys yläkoulujen oppimistuloksiin. Tulojen merkitystä ei kuitenkaan pystytty toteamaan tutkimuksen aineistolla tilastollisesti merkitseväksi regressio- tai korrelaatioanalyysien keinoin, olivat kyseessä sitten keskimääräiset tulot per henkilö tai asuntokunta. Tulokset poikkesivat aiemmin tutkitusta tässä kohdin. On esimerkiksi tutkittu, että korrelaatio oppilaaksiottoalueiden vuositulojen ja oppimistulosten välillä on Helsingissä voimakas ja tilastollisesti erittäin merkitsevä ($r = .648, p < .001$) suomenkielisten yläkoulujen kohdalla. Ruotsinkielisten koulujen kohdalla yhteys puolestaan oli Helsingissä vieläkin voimakkaampi, mutta ei tilastollisesti merkitsevä ($r = .850, p < .068$). (Bernelius, 2013a, 180 ja 188.) Oulun kohdalla vastaavia tuloksia ei tässä tutkimuksessa saatu, vaan käytettyjen alueiden keskimääräisten tulojen yhteys oli kummallakin muuttujalla mitattuna melko heikko. Näyttäisi siis siltä, että Oulussa asuinalueen tulotaso ei ole kovin voimakkaasti yhteydessä siihen, millaisia oppimistuloksia alueen yläkoulu tuottaa. Yleensä vanhempien tulojen yhteys ei ole tutkimuksissa noussut kovin merkittäväksi tekijäksi oppimistulosten ennustajana Suomessa ja kaikki tutkimukset korostavatkin nimenomaan vanhempien koulutustason merkitystä.

Toteutettujen regressioanalyysien pohjalta voidaan todeta, että yläkoulujen päättökeskiarvon vaihtelua selittää luotettavasti voimakkaimmin keskiasteen koulutuksen saaneiden osuus työvoimasta. Rakennetuista regressiomalleista lineaarisen regressiomallin edeltävyysheitojen täyttyessä tämän muuttujan selityskerroin rajaamattomilla lähialue-estimaateilla (”OULU”) oli 0,297, eli kyseisessä mallissa keskiasteen koulutuksen saaneiden osuus alueen työssäkäyvistä työvoimasta selittää 29,7 % yläkoulujen päättökeskiarvojen vaihtelusta. Vastaavasti rajattuja alueita hyödyntävässä analyysissä vastaavan mallin selityskerroin oli 0,298 eli keskiasteen koulutettujen osuus koulun lähialueella riittäisi yksistään selittämään 29,8 % yläkoulujen päättökeskiarvoista. Nämä ovat erittäin lähellä Vantaan koulujen keskiarvon vaihtelulle äidin

koulutusasteen kautta saatua selitystasetta 30 %. (Kuusela, 2010, 46–47.) Mainittakoon että hyvin lähellä olevia, mutta hiukan pienempiä tuloksia saatiin käyttämällä korkeakoulutettujen osuutta työssäkäyvistä työväestöstä selittävänä muuttujana. Vantaan ja Oulun asukasluvut ovat melko lähellä toisiaan, minkä vuoksi voisi olettaa niiden olevan paremmin vertailukelpoisia kuin Helsingin ja Oulun. Kuuselan (2010, 46–47) mukaan Vantaan koulujen keskiarvojen muita kaupunkeja pienempi selitystase selittyy äitien koulutustasojen pienemmällä vaihtelulla koulujen välillä. Vastaava selitys saattaa olla käypä myös Oulun kohdalla verrattaessa Oulua Helsinkiin ja Espooseen. Voi olla, että Oulussa yläkoulujen lähialueiden koulutustasoissa ei ole yhtä paljon vaihtelua kuin edellä mainituissa kaupungeissa. Täytyy kuitenkin huomioda, että Kuuselan mallissa tiedot kerättiin koulu-, ei aluekohtaisesti. Koulukohtaisesti kerätyillä tiedoilla päästään tarkempaan ennusteeseen, vaikka myös alueittain kerätyt tiedot ainakin pääkaupunkiseudulla vastaavat melko hyvin ennustetta oppilaiden sosioekonomisesta taustasta (Kuusela, 2010, 46).

Kun malliin lisättiin tulot, eivät mallin edeltävyys ehdot enää täyttyneet niin paljon, että mallin merkitsevyystestauksen tuloksia voitaisiin pitää pätevinä. Näin ollen mallia täytyy tarkastella hyvin kriittisesti, sillä sen ennustusvoima on melko heikko. Kun lähialue-estimaatteja ei rajattu, selityskertoimeksi mallille saatiin 0,317 ja rajatussa tilanteessa puolestaan 0,329. Lähialue-estimaatin keskiasteen koulutuksen käyneen työssäkäyvän työvoiman suhteellinen osuus ja alueen sisäinen rahatulo per henkilö pystyisivät siis yhdessä selittämään jopa yli kolmanneksen (31,7 % ja 32,9 %) yläkoulujen päättökieskiarvojen vaihtelusta. Mallin edeltävyys ehtojen rikkomuksien vuoksi mallin kahdessa eri skenaariorissa saamia tilastollisesti merkitseviä F-testin tuloksia ei voida kuitenkaan pitää suuressa arvossa, eikä mallia voida pitää kovin luotettavana ennustajana. Selitystaseen laskeminen lisättäessä malliin muuttujia koulutusasteen lisäksi

Tutkimuksen tuloksien perusteella voidaan todeta, että yläkoulujen lähialueiden asukkaiden sosioekonomista tekijöistä erityisesti koulutusaste on melko voimakkaasti yhteydessä oppimistuloksiin. Tulokset ovat pääosin hyvin linjassa aikaisempien Suomessa toteutettujen tutkimusten tulosten kanssa, vaikka pieniä eroja yhteyksien voimakkuuksissa onkin nähtävissä. Tiedetään, että koulujen oppilaiden sosioekonomiset taustatekijät – ja erityisesti vanhempien koulutustaso – ovat yleisesti ottaen voimakkaasti yhteydessä koulumenestykseen Suomessa (kts. esim. Leino ym., 2019; Kalalahti, 2012; Kuusela, 2010; Lehti & Laaninen, 2020; Salmela-Aro & Chmielewski, 2019; Vainikainen ym., 2016; Välijärvi ym., 2015). Tämän tutkimuksen perusteella Oulu ei tee poikkeusta tässä suhteessa, vaan oppilaiden sosioekonomisista

taustatekijöistä nimenomaan vanhempien koulutustaso vaikuttaisi olevan vahvasti yhteydessä koulumenestykseen. Sen sijaan vanhempien tuloilla ei vaikuttaisi Oulussa olevan kovin suurta yhteyttä koulun oppimistuloksiin, mutta oppilastasolla tilanne voi kuitenkin olla erilainen.

6 Eettiset kysymykset ja tutkimuksen luotettavuus

Tutkimuksen eettisten kysymysten kohdalla merkittävimpanä nousee esille se, että vanhemmat, joilla on kiinnostusta ja mahdollisuuksia valita lapselleen koulu, saattavat pyrkiä hyödyntämään tutkimuksen tuloksia valinnoissaan. Pitkässä juoksussa on siis mahdollista, että tutkimuksen tulokset kiihdyttävät koulujen eriytymiskehitystä. Myös tutkimuksissa on löydetty todisteita väitteelle (Bernelius, 2014, 11). Vastaavanlaisten haasteiden parissa onkin painittu aihetta käsittelevissä tai sivuavissa tutkimuksissa aiemminkin (esim. Bernelius, 2013; kts. myös Bernelius, 2014).

Edellä mainitun ongelman ratkaisu on, että tutkimuksen tulokset esitetään muodossa, josta yksittäisiä kouluja tai niiden lähialue-estimaattien tietoja ei ole mahdollista tunnistaa. Käytännössä tämä tarkoittaa sitä, että koulujen nimet jätetään pois tutkimukseen liitettävistä taulukoista ja kuvaajista, minkä lisäksi asukasmäärät, tulo- ja koulutustietoja eri alueilta ei myöskään ilmoiteta minkään alueen kohdalla erikseen. Lisäksi aineiston tunnusluvuista ja kaikista kuvaajista on piilotettu päättökeskiarvoja kuvaavien muuttujien osalta minimi- ja maksimi-arvot, sekä vaihteluväli.

Bernelius on päättänyt jättämään kartatkin julkaisematta tutkiessaan asuinalueiden ja koulujen välisiä yhteyksiä (Bernelius 2013; 2014). Tässä pro gradu -tutkielmassa päätettiin kuitenkin julkaista kartat, jotka näyttävät muodostetut lähialue-estimaatit koulujen ympärillä. Perusteluna tälle ratkaisulle on, että vaikka lähialue-estimaattien koon ja muodon avulla pysyy päättelämään jotakin alueen sosioekonomisesta taustasta, kun tieto yhdistetään vaikkapa postinumeroalueittaisiin tulotietoihin, olisi koulujen ympäristöistä kiinnostuneiden mahdollista selvittää asia myös ilman tämän tutkielman karttoja. Toisin sanottuna valvutuneet vanhemmat voisivat itsekkin tehdä johtopäätöksiä Oulun kaupungin alueita tutkimalla. Toisaalta kartat auttavat lukijaa hahmottamaan mitä tutkimuksessa oikeastaan on tehty, joten niiden julkaisemisesta on hyötyä. Tämän tutkielman karttojen käyttötarkoituksena onkin lähinnä esitellä kuinka tutkimus on toteutettu, eivätkä kartat sisällä yksityiskohtaista tietoa alueista tai kouluista.

*Reliabeeli*us tarkoittaa tutkimuksen kykyä antaa ei-sattumanvaraisia tuloksia, eli tutkimuksen tulosten toistettavuutta. Realibiteettiä arvioitaessa tarkastellaan erityisesti mittaukseen liittyviä asioita, kuten otoskokoa ja laatua, muuttujien tietojen syöttämisen huolellisuutta, sekä mahdollisia mittausvirheitä (Vilka, 2014, 150.) *Validius* puolestaan tarkoittaa tutkimuksen kykyä mitata sitä, mitä tutkimuksella yritettiin mitata. Validiteetin arvioinnissa auttavat

esimerkiksi seuraavat kysymykset: *”Miten tutkija on onnistunut teoreettisten käsitteiden operationalisoinnissa arkikielelle? – – Millaisia epätarkkuuksia mittariin sisältyy?”* (Vilkkä, 2014, 150).

Tutkimuksen realibiteettiä voidaan pitää melko hyvänä, sillä tutkimuksessa käytettävät aineistot olivat tarkkoja, niitä käsiteltiin huolellisesti ja kaikki toimenpiteet toteutettiin varmuuden vuoksi useaan kertaan, jotta voitiin varmistua siitä, että tietojen kerääminen ja syöttäminen olisivat virheettömiä. Tutkimuksen tulokset ovat täysin toistettavissa samanlaisella aineistolla ja samoja menetelmiä käyttäen. Myös aineiston käsittely pyrittiin kuvaamaan hyvin yksityiskohtaisesti. Tutkimuksessa käytettyjen otoksien voidaan katsoa edustavan perusjoukkoa melko hyvin, vaikkakin joidenkin lähialue-estimaattien kohdalla otoksen pätevyys jouduttiinkin kyseenalaistamaan. Validiutta arvioitaessa voidaan puolestaan todeta, että teoreettiset käsitteet on pyritty avaamaan selkeästi arkikielen tasolla, eikä tutkimuksessa nähdäkseen esiinny suuria systemaattisia virheitä.

Tutkimusta tehdessä ilmeni kuitenkin tekijöitä, jotka vaikuttavat tutkimuksen luotettavuuteen ja samalla sen painoarvoon. Tekijät liittyvät enimmäkseen aineistoihin, vaikuttaen hieman sekä realibiteettiin ja validiteettiin. Yleisesti ottaen tutkimuksen aineistoja voidaan pitää luotettavina, sillä niitä ovat tuottaneet Tilastokeskus, Maanmittauslaitos ja Oulun kaupunki, kaikki luotettavana pidettäviä tahoja. Tutkimuksen lähtöaineistoja voidaan siis pitää luotettavina sellaisinaan, mutta näiden aineistojen käyttökelpoisuuteen liittyvät tekijät vaikuttavat hieman tutkimuksen luotettavuuteen.

Tutkimuksen luotettavuutta heikentäviksi tekijöinä on nostettava esille, että tutkimuksen aineistot olivat peräisin eri vuosilta. Käytettävistä aineistoista yläkoulujen päättökeskiarvot ovat vuodelta 2020 ja YKR-aineisto puolestaan vuodelta 2017. Tutkimukseen laatua heikentävästi tämä vaikuttaa siten, että 2017 vuoden jälkeen Oulussa on rakennettu uusia asuinalueita. Muun muassa tästä ja tietenkin muutenkin jatkuvasta ihmisten muuttoliikkeestä seuraa se, että YKR-aineisto ei kaikilta osin ole ajan tasalla, mikä voi vääristää joidenkin alueiden poimittuja tulo- ja koulutustietoja. Voidaan kuitenkin todeta, että pääosin Oulun kaupunki ei ole vuosien 2017 ja 2020 välillä muuttunut niin paljon, että tämä aiheuttaisi kovin merkittävää ongelmaa tutkielman kannalta. Tuloksia tarkastellessa on kuitenkin otettava huomioon, että kyseisellä asialla on varmasti ainakin jonkinlainen vaikutus muun muassa havaintojen arvoihin.

Lisäksi on otettava huomioon, että aineistosta on poistettu kouluja. Koulujen poistamisen perusteluita on käsitelty jo aiemmin. Oulun kouluverkossa mukana olevat yksityiset koulut, sekä kaupungin oma erityiskoulu muokkaavat todennäköisesti jonkin verran koulujen todellisia oppilaaksiottoalueita ja/tai koulujen oppilaspohjia. Oulun kaupungin käyttämät suuret oppilasalueet aiheuttavat myös sen, että lähialue-estimaatit ovat nimensä mukaisesti vain estimaatteja - oletuksia. Todellisuudessa oppilaat saattavat käydä oman alueensa toista yläkoulua, tai jopa toisen alueen koulua esimerkiksi hakeuduttuaan painotettuun opetukseen. Tutkimuksessa oltiin kuitenkin kiinnostuneita siitä, selittävätkö yläkoulun lähialueen sosioekonomiset tekijät kyseisen koulun oppimistuloksia. Näin ollen oppilaiden valikoitumisella muuhun kuin lähimpään yläkouluun ei ole tutkimuksen tulosten kannalta merkitystä, mutta tuloksia ja niiden taustalla vaikuttavia tekijöitä tulkitessa edellä mainitut asiat on otettava huomioon. Koulujen oppilaiden sosioekonomiset taustat eivät nimittäin edellä mainituista syistä välttämättä vastaa koulun lähialueen sosioekonomista taustaa.

Esille on nostettava myös YKR-aineistoissa esiintyvät tietosuojaan liittyvät tekijät. Käytettävissä olleen aineiston tietoja on salattu tiettyssä ruutujen väestömäärien perusteilla. Tietoja on esimerkiksi salattu merkitsemällä esimerkiksi tulotietojen kohdalle -1, jos ruudun asukasmäärä on alle 10. Tämä johtaa siihen, että aineistossa painottuvat tiheästi asutut ruudut, jouduttaessa jättämään pois niitä ruutuja, joissa asukkaiden määrä jää alle salausrajan. Käytännössä harvaanasutuilla alueilla suurin osa ruuduista jäi pois; esimerkiksi erään lähialue-estimaatin ruuduista pystyttiin poimimaan analyysissä hyödynnettäväksi vain noin 2 % alueen ruuduista.

Edellä kuvattu vääristää alueellisten tietojen keräämistä ja muuttaa lähes kaiken datan koskemaan jonkinlaista keskustaa tai muuta asumisen tihentymää. Tämä on ongelmallista erityisesti sen vuoksi, että tutkimuksesta jäävät pois harvaanasutulle seudulle ominaiset piirteet. On esimerkiksi mahdollista, että harvaanasutuilla alueilla asuu suhteessa enemmän matalamman koulutuksen saaneita ihmisiä kuin kaupunkialueilla, kyläkeskitymissä tai muissa taajamissa. Näin ollen suurten alueiden, joiden sisään mahtuu kumpiakin asumismuotoja, kerätyt tilastotiedot saattavat vääristää esimerkiksi korkeakoulutetun työväestön osuutta estimaatin työväestöstä ja hilata sen suhteellista osuutta ylöspäin. Harvaanasuttujen alueiden sosioekonomisista piirteistä on siis vaikea saada todellisuutta vastaavaa tietoa. Harvaanasuttujen seutujen sosioekonominen rakenteen yhteys oppimistuloksiin jää tietosuojaominaisuuksien vuoksi näin ollen todellisuudessa käsittelemättä, vaikka harvaanasutut alueet ovatkin näennäisesti mukana osassa analyyyseistä.

Toisaalta on yleisesti tiedossa myös, että koulujen oppimistuloksien arviointi ei ole kovin vahvasti kontrolloitua ja koulujen välillä saattaa olla suuriakin eroja siinä, millaisilla tiedoilla minkäkin arvosanan saa (kts. esim. Ouakrim-Soivio, 2013). Päätöskeskisarvojen käyttäminen koulujen välisten osaamiserojen arvioinnissa on siis hieman kyseenalaista. Oppimistulosten arvioinnissa päätöskeskisarvoja parempia olisivatkin olleet standardoidut testit. Olemassa olevien testitulosten puuttuessa päädyttiin kuitenkin käyttämään apuna päätöskeskisarvoja tiedostaen kuitenkin tähän liittyvät haasteet.

Tutkimuksen *kokonaisluotettavuutta*, joka muodostuu validiudesta ja reliaabeliuksesta voidaan Vilkan (2014, 152) mukaan parantaa pitämällä tutkimusprosessin aikana huolta seuraavista asioista: Tutkitaan sitä, mitä pitikin tutkia – asia- ja tutkimusongelmat on määritelty selkeästi. Tutkimuksen perusjoukon valinta perustellaan hyvin. Otosjoukko määritellään huolellisesti ja se sisältää kaikki perusjoukon ominaisuudet, sekä on kooltaan riittävä. Valitaan sopiva otantamenetelmä ja aineiston keräämisen tapa, sekä analyysimenetelmät, joilla voidaan saada merkittävää tietoa tutkittavasta asiasta. Varmistetaan, että tarvittava tieto on mukana mittauksessa. Tiedot syötetään ja tarkistetaan huolellisesti ennen tallentamista, sekä toteutetaan prosessi muutenkin huolellisesti. (Vilka, 2014, 152.) Määrällisen tutkimuksen kokonaisluotettavuuteen vaikuttavat myös tieteelliselle tutkimukselle yleisesti asetetut kriteerit, kuten esimerkiksi se, että tutkimuskohde on täsmällisesti määritetty, tutkimus tuottaa jotain uutta, tutkimus on avoin, tulokset ovat puolueettomia ja vääristelemättömiä, tutkimus on hyödyllinen ja sen tieto on toisten sovellettavissa ja tutkimus noudattaa hyvää tieteellistä etiikkaa ja säilyttää tutkittavien yksityisyyden, sekä ammatti- ja liikesalaisuuden. Lisäksi tutkimuksen tulisi olla ajankohtainen ja sen olisi annettava riittävät perusteet esimerkiksi julkiselle keskustelulle. (Vilka, 2014, 154.)

Tutkimuksen luotettavuutta heikentävät yksityiskohdat liittyvät pääosin joihinkin aineiston ominaisuuksiin. Vaikka aineistoihin liittyikin tutkimuksessa kuvattuja ongelmia, on tutkimus kuitenkin pyritty toteuttamaan hyvää tieteellistä harkintaa käyttäen, tuoden esille havaitut epäkohdat. Jo pelkästään epäkohtien esille nostamisen voidaan ajatella nostavan tutkimuksen luotettavuutta. Tutkimuksen kokonaisluotettavuutta voidaankin nähdäkseni pitää melko hyvänä, sillä esiin nousseet luotettavuutta heikentävät tekijät eivät todennäköisesti vaikuta merkittävästi saatuihin tuloksiin ja ne on otettu tulosten tulkinnassa huomioon. Tutkimusaihe on lisäksi jo valmiiksi ajankohtainen ja tuottaa uutta tietoa laajentamalla koulujen eriytymisen tutkimusta Pohjois-Suomeen. Lisäksi tutkimuksessa käytetyt menetelmät olivat sellaisia, että niitä on aiemmissa asiaan liittyvissä tutkimuksissa hyödynnetty tavalla tai toisella. Tutkimusta

tehtäessä hyödynnettiin apuna aiemmin Suomessa tehtyjä tutkimuksia ja pyrittiin käyttämään samoja menetelmiä, jotta tutkimusten tulokset olisivat edes jollain tasolla vertailukelpoisia. Tältäkin osin tutkimuksen luotettavuutta voidaan pitää melko hyvänä ja tuloksia vertailukelpoisina muihin tutkimuksiin nähden.

On huomioitava, että tutkimuksen tulokset eivät kuitenkaan olekaan suoraan yleistettävissä. Tästä huolimatta niistä voidaan havaita yhtäläisyyksiä aiempien tutkimusten tuloksiin. Esimerkiksi vanhempien koulutustaustan voimakas yhteys lapsen oppimistuloksiin oli aiemmissa tutkimuksissa todettu (kts. esim. Lehti & Laaninen, 2020; Salmela-Aro & Chmielewski, 2019). Voidaan todeta, että tutkimuksen tulokset eivät ole ristiriidassa aiempien tutkimuksien tuottaman tiedon kanssa, vaan sen sijaan tukevat näiden tutkimuksien tuloksia. Useiden aiempien tutkimusten ollessa jokseenkin samoilla linjoilla tuloksien suhteen, voidaan olla sitä mieltä, että tämän tutkimuksen luotettavuutta voidaan pitää ainakin kohtuullisen hyvänä, ja että tutkimusta voidaan pitää tämänkin puolesta vähintään suuntaa antavana perustutkimuksena.

7 Pohdintaa

Tutkimuksen tulokset osoittivat, että Oulussa on nähtävissä yläkoulujen eriytymistä, jota voidaan selittää niiden lähialueiden sosioekonomisilla tekijöillä. Tulokset olivat melko hyvin linjassa myös aiemman tutkimustiedon kanssa, vaikka yhteydet asukkaiden koulutustasojen ja alueiden yläkoulujen kanssa jäivätkin heikommiksi, kuin muualla Suomessa tehdyissä tutkimuksissa. Tutkimuksessa pystyttiin vastaamaan asetettuihin kysymyksiin ja todettiin, että myös Oulussa yläkoulujen oppimistulokset vaikuttaisivat olevan yhteydessä koulujen lähialueiden sosioekonomisiin tekijöihin. Tulokset olivat hyvin samansuuntaisia kuin aiemmissa vastaavia yhteyksiä selvittävässä tutkimuksessa on saatu: Suomessa alueen asukkaiden koulutustaso selittää parhaiten koulujen oppimistuloksia, tulotasolla sen sijaan ei ole yhtä suurta merkitystä.

Tulojen merkityksen pienuuteen voi olla useita syitä. Voi olla, että Oulussa niin sanotut ”hyvät koulut” eivät välttämättä sijaitse ”hyvillä alueilla”. Tämä voisi tapahtua esimerkiksi niin, että oppilaita hakeutuu paljon muuhun kuin omaan lähikouluunsa esimerkiksi painotetun opetuksen perässä, mutta vanhemmat eivät ole pyrkineet muuttamaan lähelle toivottua koulua, eli asuntomarkkinat eivät ole reagoineet. Toisaalta voi olla myös niin, että oppilaita sijoitellaan enimmäkseen muihin kuin lähimpään yläkouluun. Tällöin koulun lähialueen asukkaiden tulot eivät vastaisi oppilaiden vanhempien tuloja. Koulun oppilaspohjan ja alueen asukkaiden sosioekonomiset profiilit siis poikkeaisivat toisistaan. Tämä vaikuttaa kuitenkin epätodennäköiseltä vaihtoehdolta. Myös tuloja kuvaavissa muuttujissa voi olla häiriöitä; koulun lähialueella saattaa esimerkiksi asua paljon hyvätuloisia, joilla ei ole lapsia, ja oppilaat tulisivat alueella vähemmistönä olevista pienituloisista perheistä. Tällöin alueen tuloja kuvaavat muuttujat saattaisivat näyttää alueen tulotason olevan korkea, mutta oppilaat tulisivat kuitenkin matalatuloisista perheistä samalta alueelta. Tässä tapauksessa oppilaspohjan sosioekonominen profiili ei vastaisi alueen yleiskuvaa, mikä saattaisi näkyä heikkona korrelaationa koulun tulosten ja lähialueen laskennallisen tulojen välillä.

Hieman aiemmista tutkimuksista poikkeavaa oli, että yhteys yläkoulujen päättökeskiarvoihin näkyi kaikista voimakkaimmin keskikasteen koulutuksen saaneiden osuutta alueella kuvaavan muuttujan kautta ja negatiivisesti. Miksi yhteys juuri tämän muuttujan kohdalla oli kaikista voimakkain? Aiemmissa tutkimuksissa yhteyttä on käsitelty korkeakoulutettujen ja vain perusasteen koulutuksen saaneiden osuuksien kautta, joten tutkimuksen tulos oli hieman yllätys. Myös negatiivinen yhteys on pohtimisen arvoinen asia. Minkä vuoksi keskiasteen

koulutuksen saaneen työssäkäyvän työvoiman osuuden kasvaessa oppimistulokset heikkenevät?

Yksi selitys voisi olla se, että - kuten tiedetään - keskiasteen käyneiden osuuden kasvaessa korkeakoulutettujen osuus laskee lähes täydellisessä lineaarisessa riippuvuussuhteessa. Päätökeskiarvojen laskemista tämä voisi selittää sitä kautta, että korkeakoulun käyneillä vanhemmilla olisi jotain sellaisia resursseja, jotka edesauttavat lasten koulumenestystä. Tämä selitys kävisi hyvin yksiin aiemmin tutkitun kanssa. Toisaalta asiaa voidaan lähestyä myös siltä kannalta, että korkeimmillaan keskiasteen koulutuksen saaneilla olisi joitain sellaisia resursseja, jotka haittaavat koulussa menestymistä. Tällaisten resurssien keksiminen on kuitenkin hieman haastavaa ja ainoa järkevä tähän kategoriaan mahtuva selitys voisi olla, että tässä koulutusluokassa asenteet eivät tue oppimista. Kuulostaisi kuitenkin erikoiselta, että perinteisesti koulutusta hyvin paljon arvostavassa suomalaisessa yhteiskunnassa korkeimmillaan keskiasteen koulutuksen käyneiden asenne olisi koko ryhmän tasolla sellainen, että se vaikuttaa tutkimuksen tuloksiin. Näiden pohdintojen puolesta olen itse taipuvainen selittämään asian niin, että yhteys tapahtuu nimenomaan korkeakoulutettujen osuuden laskemisen kautta, ja että jokin muu syy selittää asiaa paremmin. Keskiasteen koulutetuilla saattaa esimerkiksi olla enemmän lapsia tai heidän lapsensa saattavat olla useammin lähikoulussa, sillä tutkimukset ovat osoittaneet korkeakoulutettujen osallistuvan kouluvalintamarkkinoille ahkerammin ja menestyksekkäämmin (esim. Kalalahti ym., 2015a; Kalalahti ym., 2015b).

Kouluvalinnat ovat asia, jota olisi järkevää tutkia myös Oulussa eri näkökulmista, jotta saataisiin hyvä käsitys niiden vaikutuksista esimerkiksi koulujen eriytymiseen. Kouluvalintojen ja valintamarkkinoiden tutkiminen saattaisi antaa myös muuta tärkeää tietoa Oulun kouluista; hyljeksitäänkö esimerkiksi joitain kouluja tai niiden alueita kuten Turussa, tai ovatko jotkin koulut poikkeuksellisen vetovoimaisia syystä tai toisesta? Tällainen tutkimus antaisi lisäperspektiiviä esimerkiksi tämän pro gradu -tutkielman tulosten tulkintaan.

Kuten jo tutkimuksen taustateoriaa esitellessäni toin ilmi, hyvien ja huonojen koulujen eriytymiskehitys tarttunee pikkuhiljaa myös keskitason kouluihin. Tämä on ongelmallista, sillä keskitason koulujen alkaessa eroamaan toisistaan merkittävässä määrin eriytymiskehitys siirtyy koskettamaan yhä suurempaa osaa suomalaisista perheistä. Tämän tapahtuessa koulujen eriytymiskehityksellä alkaa todennäköisesti olla myös yhä vahvempia tasa-arvoon – eikä pelkästään koulutukselliseen – vaikuttavia seurauksia. Esimerkiksi alueellinen epätasa-

arvoisuus tulee todennäköisesti kasvamaan parempien koulujen houkutellessa lähialueelleen sosioekonomisesti paremmin pärjääviä perheitä. Näin hyväosaisuus kasautunee tietyille alueille, huono-osaisuuden kasautuessa toisille. Muitakin laajoja yhteiskunnallisia ongelmia saattaa syntyä kehityksen tuloksena.

Uhkana on koulutuksellisen tasa-arvon mureneminen kehityksen seurauksena, kun koululla on yhä merkittävämpi rooli oppilaan koulutustuloksissa ja oppimisessa. Ongelmallista heikkojen ja hyvien koulujen eriytymisessä on myös se, että eriytymiskehitys tarttunee pikkuhiljaa myös keskitason kouluihin. Pienetkin erot voidaan toki nähdä koulutuksellista tasa-arvoa heikentävinä, mutta totuus lienee kuitenkin se, että eroja syntyy aina niin kauan kuin ihmiset eivät ole toistensa kopioita. Suuret erot kuitenkin kielinevät jostain perustavaa laatua olevasta rakenteellisesta ongelmasta suomalaisessa koulujärjestelmässä, minkä vuoksi erojen syyt on parasta pyrkiä jäljittämään mahdollisimman nopeasti, sillä kuvaamani kaltainen kehityskulku on mielestäni ehdottomasti epätoivottavaa, ja sen voidaan katsoa uhkaavan jopa koko Suomen mittapuulla yksilöiden yhteiskunnallista tasa-arvoa, johon kansakunta niin kovasti pyrkii, ja jonka pitkälle edenneestä toteutuksesta se perinteisesti on ylpeä.

Eriytymiskehitykseen liittyen; mielenkiintoista oli, että tutkimustuloksissa havaittiin eroja sivistys- ja kulttuurialueiden avulla muodostettujen lähialue-estimaattien ja oppilasaluerajaamattomien tulosten välillä. Esimerkiksi korrelaatio päättökeskiarvojen ja keskiasteen koulutuksen saaneen työvoiman suhteellisen osuuden välillä oli voimakkaammin negatiivinen SIKU-alueiden avulla muodostettujen lähialue-estimaattien kohdalla verrattuna rajoja huomioimattomaan malliin (vrt. esim. $r_s = -,547$, $p = ,013$ vs. $r_s = -,564$, $p = ,010$). Mielenkiintoisesti myös korrelaatioiden tilastolliset merkitsevyydet kasvoivat kautta linjan vertailtaessa rajaamattomia ja SIKU-rajattuja lähialue-estimaatteja. Nämä erot merkitsevyyksissä olivat kuitenkin hyvin pieniä. Eroja syntyi sekä poikkeavien havaintojen ollessa mukana analyysissä, että silloin kun nämä havainnot oli poistettu aineistosta. Erot myöskään korrelaatioissa eivät olleet järin suuria, mutta kuitenkin olemassa olevia. On vaikea arvioida mistä havainnot kertovat, tai kertovatko mistään, eikä siihen myöskään haettu vastauksia tässä tutkimuksessa. Asia on kuitenkin mielenkiintoinen merkilläpantava. Voisiko olla niin, että Oulun kaupungin oppilasaluejaossaan käyttämät sivistys- ja kulttuurialueiden rajat saattaisivat jopa vahvistaa koulujen eriytymistä?

Tutkimus oli ensimmäinen laatuaan Oulussa, viitoittaen samalla tietä muille vastaaville muualla kuin pääkaupunkiseudulla, Turussa tai muualla Etelä-Suomessa tehtäville

tutkimuksille. Tutkimuksen ansioksi voitaneenkin laskea koulujen eriytymiseen liittyvän aluetutkimuksen kentän avaaminen Oulussa. Tutkimuksessa oltiin kiinnostuneita myös siitä, poikkeavatko pohjoisessa Suomessa saadut tulokset eteläisen Suomen vastaavista. Aineiston pienen koon ja joidenkin tutkimuksen luotettavuutta hieman heikentävien tekijöiden vuoksi ei voida kuitenkaan olla aivan varmoja siitä, poikkeavatko Oulussa saadut tulokset todellisuudessa merkittävästi pääkaupunkiseudusta, mutta viitteitä vertailukohteita heikommista oppimistulosten ja alueellisten sosioekonomisten tekijöiden yhteyksistä tuloksista on kuitenkin tulkittavissa. Jos oletetaan eron olevan todellinen, eli että se ei johdu esimerkiksi aineiston pienestä koosta, on mielenkiintoista pohtia, miksi yhteydet selittävän ja selitettävien muuttujien välillä Oulussa ovat Helsinkiä pienempiä. Todennäköisesti syy löytyy kaupunkien erilaisista ominaisuuksista, sillä pääkaupunkiseudun voidaan katsoa poikkeavan vahvasti muusta Suomesta monellakin tapaa.

Kaupungit luonnollisesti poikkeavat toisistaan paljon jo asukasmääriltään. Voi olla esimerkiksi niin, että Helsingin Oulua merkittävästi suuremmat asukasmäärät johtavat voimakkaampaan alueelliseen eriytymiseen, joka sitten heijastuu myös kouluihin. Esimerkiksi Vilkaman ja Hirvosen (2018) mukaan on myös tutkittu, että suurimmissa kaupungeissa polarisaatio hyvä- ja huono-osaisten välillä näyttäytyy tyypillisesti muuta maata jyrkempänä. Helsinki ja Oulu kumpikin voidaan nähdä suomalaisessa kontekstissa suurina kaupunkina, mutta Helsingin asukasmäärä on kuitenkin noin kolminkertainen verrattuna Ouluun (vrt. 653 835 vs. 205 489 vuonna 2019 (Kuntaliitto, 2019)). Tämä on mielestäni melko suuri ero ja Oulua olisikin kenties viisaampaa verrata Vantaaseen, joka on asukasluvultaan (233 775 vuonna 2019 (Kuntaliitto, 2019)) melko lähellä Oulua. Vantaalla koulujen toimintaedellytykset ovatkin alueellisesta näkökulmasta suhteellisen tasaiset verrattuna muihin pääkaupunkiseudun kuntiin nähden (Bernelius, 2015, 638 ja 640). Voikin olla niin, että Oulu vertautuu melko hyvin Vantaaseen suhteessa pääkaupunkiseutuun, vaikka on toki otettava huomioon, että Vantaa sijaitsee aivan Helsingin kupeessa, mikä vaikuttaa varmasti tavalla tai toisella sen kehitykseen.

Voi olla myös, että Oulun koulutuspolitiikka on onnistuneempaa kuin vaikkapa Helsingissä. Voidaan pohtia sitä, onko näiden kaupunkien kouluvalintamarkkinoissa jotain sellaisia eroja, jotka selittävät oppimistulosten ja alueellisten sosioekonomisten tekijöiden korrelaatioiden väliset erot. Saattaa olla esimerkiksi niin, että Helsingissä kouluvalintaa tehdään enemmän kuin Oulussa, mikä johtaa koulujen oppimistulosten välisten erojen kasvuun. Toisaalta alueellinen segregaatiokin voi olla voimakkaampaa pääkaupunkiseudulla kuin Oulussa. Tässäkin suhteessa vaikkapa Vantaa voisi olla parempi vertailukohde Oululle, mutta Vantaalla ei

kuitenkaan ole käsittäkseni tehty vastaavaa tutkimusta oppimistuloksiin ja alueisiin liittyen. Kouluvalintamarkkinat Helsingin ja Oulun välillä poikkeavat todennäköisesti vahvasti, sillä Helsingissä on huomattavasti enemmän yläkouluja (kts. esim. Helsingin kaupunki, 2020) kuin asukasmäärältään pienemmässä Oulussa, ja näin ollen todennäköisesti myös enemmän valinnanvaraa koulujen suhteen. Kouluvalintamarkkinat ja niiden ominaisuudet voivat myös selittää osaltaan tuloksien eroja.

Oulu on myös hyvin laajalle levittäytynyt kaupunki, jossa kaupunkikeskustan osuus on suhteellisen pieni. Kouluverkostoon kuuluvien koulujen etäisyys voi myös selittää sitä, miksi korrelaatiot Oulussa eivät olleet yhtä suuria kuin Helsingissä. Koulujen välisten matkojen kasvaessa niiden lähialue-estimaattien sisälle mahtuu suuremmalla todennäköisyydellä myös toisistaan sosioekonomisesti poikkeavia alueita. Hyvin suurten lähialue-estimaattien sisään voi mahtua esimerkiksi korkeasti koulutettujen ja hyvätuloisten alueita, sekä vähän koulutettujen ja vähätuloisten alueita. Näiden osuessa saman lähialue-estimaatin sisään, estimaatin ominaisuudet siirtyvät kohti näiden keskiarvoa. Jos alueet ovat riittävän suuria, voidaan arvioida, että niiden kaikkien sosioekonomiset keskiarvo-ominaisuudet olisivat lähes samat. Näin ollen pienemmät etäisyydet koulujen välissä, ja sitä kautta pienemmät lähialue-estimaatit kuvaavat todennäköisesti hieman paremmin alueiden todellisia sosioekonomisia taustamuuttujia.

Koulujen ja alueellisten ominaisuuksien välinen tutkimus keskittyy mielestäni harmillisen paljon Etelä-Suomeen, pääkaupunkiseudulle ja erityisesti Helsinkiin. Tällä hetkellä vaikuttaa siltä, että tutkimusta tekee vain korkeintaan muutama henkilö ja heidän tutkimuksensa keskittyvät tietyille alueille. Tilausta koulujen tutkimukselle mielestäni on edelleen, ja tutkimuksellisen tyhjiön täyttämiseksi sitä kannattaisi tehdä muuallakin kuin pääkaupunkiseudulla tai Turussa. Suomessa on kuitenkin lukuisia erilaisia kaupunkeja, joista olisi mielenkiintoista saada vertailukelpoisia tuloksia. Esimerkiksi vaikkapa Rovaniemi, Kuopio ja Vaasa. Jo näitä kaupunkeja tutkimalla saataisiin ympäri Suomea tuloksia. Lisätutkimusta tarvitaan, ja toivottavasti koulujen ja niiden lähialueiden tutkimuskenttä/-perinne laajenee Suomessa myös nykyisen ydinalueensa ulkopuolelle.

Lähteet

- Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, 18(3), 91–93. <https://doi.org/10.1016/j.tjem.2018.08.001>
- Albrecht, J. (2007). *Key concepts and techniques in GIS*. Los Angeles: SAGE Publications.
- Aluehallintavirasto. (2016). Toimialueet. *Avi.fi*. Haettu 13.11.2020 osoitteesta <https://www.avi.fi/web/avi/toiminta-alue>
- Antikainen, A., Rinne, R., & Koski, L. (2013). *Kasvatussosiologia* (5. uud. p.). Jyväskylä: PS-kustannus.
- Bernelius, V. (2005). *Onko oppimistulokset valettu betoniin?: Tutkimus Helsingin kaupunkirakenteen ja peruskoulujen oppimistulosten välisestä yhteydestä ja kouluvalintojen vaikutuksista*. (Pro Gradu). Helsinki: Helsingin yliopisto. <http://urn.fi/URN:NBN:fi-fe200801301053>
- Bernelius, V. (2010). Alueellinen eriytyminen heijastuu kouluihin. Teoksessa M. Rimpelä, & V. Bernelius (toim.) *Peruskoulujen oppimistulokset ja oppilaiden hyvinvointi eriytyvällä Helsingin seudulla MetrOP-tutkimus 2010–2013. Mitä tiedettiin tutkimuksen käynnistyessä keväällä 2010?* (19–23). Helsinki: Helsingin yliopisto.
- Bernelius, V. (2011). Osoitteenmukaisia oppimistuloksia? Kaupunkikoulujen eriytymisen vaikutus peruskoululaisten oppimistuloksiin Helsingissä. *Yhteiskuntapolitiikka*, 76(5), 479–493. <http://urn.fi/URN:NBN:fi-fe201209117868>
- Bernelius, V. (2013a). *Eriytyvät kaupunkikoulut: Helsingin peruskoulujen oppilaspohjan erot, perheiden kouluvalinnat ja oppimistuloksiin liittyvät aluevaikutukset osana kaupungin eriytymiskehitystä*. Väitöskirja. Helsinki: Helsingin yliopisto
- Bernelius, V. (2013b). Koululaiset kaupungissa: Kouluvalintoja ohjaavat naapurustojen piirteet ja valintojen vaikutus koulujen eriytymiseen Helsingissä. *Yhdyskuntasuunnittelu*, 51(1), 8–27. Haettu 13.11.2020 osoitteesta: <https://www.yss.fi/journal/koululaiset-kaupungissa/>

- Bernelius, V. (2014). Koulut kartalle - ilman karttaa. *Positio: paikkatiedon erikoislehti*, 2014(1), 10-12. Haettu 13.11.2020 osoitteesta http://www.paikkatietoikkuna.fi/web/fi/positio_1_2014_koulut_kartalle_ilman_karttaa
- Bernelius, V. (2015). Pääkaupunkiseudun koulujen naapurustot – missä erot kasvavat? *Yhteiskuntapolitiikka*, 80(6), 635–642. <http://urn.fi/URN:NBN:fi-fe2015121023629>
- Bernelius, V. (8.3.2019). ”Kansainväliset koulut” ja ”maahanmuuttajakoulut” – kaupunkisegregaatio, etnisyys ja mielikuvien kartat. *Versus*. Haettu osoitteesta <https://www.versuslehti.fi/kriittinen-tila/kansainvaliset-koulut-ja-maahanmuuttajakoulut-kaupunkisegregaatio-etnisyys-ja-mielikuvien-kartat/>
- Chmielewski, A. K. (2019). The Global Increase in The Socioeconomic Achievement Gap, 1964 to 2015. *American Sociological Review*, 84(3), 517–544 <https://doi.org/10.1177/0003122419847165>
- Dhalmann, H., Vaattovaara, M., & Vilkama, K. (2013). Hyvää kasvuympäristöä etsimässä: Asuinalueen ja koulun merkitys lapsiperheiden muuttopäätöksille pääkaupunkiseudulla. *Yhdyskuntasuunnittelu*, 51(4), 11-29.
- Elo, P. (31.3.2018). Lukijalta: Koulujen eriytyminen vakava uhka. Haettu osoitteesta <https://www.ts.fi/lukijoilta/3894144/Lukijalta+Koulujen+eriytyminen+vakava+uhka>
- Eskonen, H., de Fresnes, T. & Pietarinen, E. (16.6.2020). Peruskoulun pitäisi taata samat lähtökohdat köyhien ja varakkaiden lapsille, mutta repikö korona tasa-arvon kappaleiksi? *Yle.fi*. Haettu osoitteesta <https://yle.fi/uutiset/3-11394459>
- From, T., Kalalahti, M., Mietola, R., Paakkari, A., Sahlström, F., Varjo, J., & Vartiainen, H. (2014). Eriytyvä peruskoulu. *Yhteiskuntapolitiikka*, 79(5), 635–642 <http://urn.fi/URN:NBN:fi-fe201502091550>
- Harju-Luukkainen, H., Sulkunen, S., & Vettenranta, J. (2015). Enemmän yhtäläisyyksiä kuin eroja. Ruotsinkieliset pienet ja suuret koulut PISA 2009 -tulosten valossa. *Nuorisotutkimus*, 33(3–4), 59–73.

- Hautamäki, J., Kupiainen, S., Marjanen, J., Vainikainen, M-P., & Hotulainen, R. (2013). *Oppimaan oppiminen peruskoulun päättövaiheessa: Tilanne vuonna 2012 ja muutos vuodesta 2001*. Helsinki: Helsingin yliopisto.
- Heikkilä, T. (2014). *Tilastollinen tutkimus*. (9. uud. p.). Helsinki: Edita.
- Helsingin kaupunki. (2020). Helsingin peruskoulut. Haettu 13.11.2020 osoitteesta <https://www.hel.fi/peruskoulut/fi/koulut/>
- Honkanen, P. (2018). *Esimerkki ja simulointilaskelmia sosiaaliturvan muutoksista 2012–2019*. Helsinki: SOSTE Suomen sosiaali- ja terveys ry. Haettu 13.11.2020 osoitteesta https://www.soste.fi/wp-content/uploads/2018/11/soste_sosiaaliturvaraportti_14022018_valmis.pdf
- Johansson, E. (20.8.2019). "Keskiluokkaiset vanhemmat voivat kouluvalinnoilla tuottaa pelkäämäänsä kehitystä", sanoo tutkija – se, mihin kouluun lapsesi laitot, voi aiheuttaa eriarvoisuutta. *Yle.fi*. Haettu osoitteesta <https://yle.fi/uutiset/3-10928059>
- Jokivuori, P., & Hietala, R. (2014). *Määrällisiä tarinoita: Monimuuttujamenetelmien käyttö ja tulkinta*. Helsinki: Docendo.
- Kalalahti, M. (2012). Perhetaustan vaikutus tyttöjen ja poikien koulunkäyntiin. *Kasvatus*, 43(4), 375–390.
- Kalalahti, M. (2014). *Muuttuvat koulutusmahdollisuudet: Nuorten sosiaaliset hierarkiat ja koulumenestys*. (Väitöskirja). Helsinki: Helsingin yliopisto. <http://urn.fi/URN:ISBN:978-952-10-9118-6>
- Kalalahti, M., Silvennoinen, H., & Varjo, J. (2015a). Kouluvalinnat kykyjen mukaan? Erot painotettuun opetukseen valikoitumisessa. *Kasvatus*, 46(1), 19–35.
- Kalalahti, M., Silvennoinen, H., & Varjo, J. (2015b). Luokittunut kouluvalinta ja mahdollisuuksien tasa-arvo. Teoksessa P. Seppänen, M. Kalalahti, R. Rinne & H. Simola (toim.) *Lohkoutuva peruskoulu: Perheiden kouluvalinnat, yhteiskuntaluokat ja koulutuspolitiikka* (371–400). Jyväskylä: Suomen kasvatustieteellinen seura.

- Kalalahti, M., & Varjo, J. (2016). Lähikoulupolut ja painotetut valinnat. Teoksessa H. Silvennoinen, M. Kalalahti & J. Varjo (toim.) *Koulutuksen tasa-arvon muuttuvat merkitykset* (37–68). Jyväskylä: Suomen kasvatustieteellinen seura.
- Kalenius, A. (2014). *Koulutus, työllisyys ja työttömyys*. Helsinki: Opetus- ja kulttuuriministeriö. <http://urn.fi/URN:ISBN:978-952-263-291-3>
- Kauppinen, P. (18.3.2020). Pääkirjoitus: Koronakotikoulu sysää lapset yhä suurempaan epätasa-arvoon. *Iltalehti.fi* Haettu osoitteesta <https://www.iltalehti.fi/paakirjoitus/a/85c8cd5f-8257-4bf8-a976-a70a809afac9>
- Kärkkäinen, T. (2004). Koulutuksen ja lapsi-vanhempisuhteen yhteys elämässä selviytymiseen: Sosiaalinen perimä ja koulutuskulttuurisen pääoman periytyminen sukupolvesta toiseen Väitöskirja. Helsinki: Helsingin yliopisto. <http://urn.fi/URN:ISBN:952-10-1292-7>
- Kivinen, O., Hedman, J., & Kaipainen, P. (2012). Koulutusmahdollisuuksien yhdenvertaisuus suomessa. eriarvoisuuden uudet ja vanhat muodot. *Yhteiskuntapolitiikka*, 77(5), (19–35). <http://urn.fi/URN:NBN:fi-fe201301021030>
- Koivisto, A. (2018). *Koulutustason periytyminen ja koulutuksellinen tasa-arvo Suomessa*. (Kandidaatintyö). Oulu: Oulun yliopisto. <http://urn.fi/URN:NBN:fi:oulu-201809132782>
- Komulainen, S. (2012). ”White Flight” in Finland? A Qualitative Study into Finnish-born Families’ Housing and School Choices in Turku. *Finnish Yearbook of Population Research*, 47, 51–64. <https://doi.org/10.23979/fypr.45074>
- Koskela, A. (2016). Koulutustason periytyvyys Suomessa – Miksi maisterin lapsesta tulee maisteri? Teoksessa K. Söder, & A. Karlsson (toim.) *Suomen koulutuspolitiikan tulevaisuus*. (118–124). Helsinki: Into.
- Kosunen, S., Seppänen, P., & Rinne, R. (2015). Yläluokka ja kouluvalinta kaupungeissa. Teoksessa P. Seppänen, M. Kalalahti, R. Rinne & H. Simola (toim.) *Lohkoutuva peruskoulu: Perheiden kouluvalinnat, yhteiskuntaluokat ja koulutuspolitiikka*. (469–508). Jyväskylä: Suomen kasvatustieteellinen seura.

- Kosunen, S., Bernelius, V., Seppänen, P., & Porkka, M. (2020). School Choice to Lower Secondary Schools and Mechanisms of Segregation in Urban Finland. *Urban Education*, 55(10), 1461–1488. <https://doi.org/10.1177/0042085916666933>
- Kosunen, S., Carrasco, A., & Flores, C. (2015). Kosunen, S., Carrasco, A., & Tironi, M. (2015). The Role of 'Hot' and 'Cold' Knowledge in the Choice of Schools in Chilean and Finnish Cities. Teoksessa P. Seppänen, A. Carrasco, M. Kalalahti, R. Rinne, & H. Simola (toim.) *Contrasting Dynamics in Education Politics of Extremes: School Choice in Chile and Finland* (139–157). Rotterdam: Sense Publishers.
- Kuivalainen, S. (24.3.2020). Korona ei kohtelee lapsia samalla tavalla. *Eläketurvakeskus*. Haettu osoitteesta <https://www.etk.fi/blogit/korona-ei-kohtelee-lapsia-samalla-tavalla/>
- Kuntaliitto. (3.4.2019). Kaupunkien ja kuntien lukumäärät ja väestötiedot. Haettu osoitteesta <https://www.kuntaliitto.fi/tilastot-ja-julkaisut/kaupunkien-ja-kuntien-lukumaarat-ja-vaestotiedot>
- Kuusela, J. (2010). Oppilaiden sosioekonomisen taustan yhteys koulumenestykseen koulutasolla. Teoksessa M. Rimpelä, & V. Bernelius (toim.) *Peruskoulujen oppimistulokset ja oppilaiden hyvinvointi eriytyvällä Helsingin seudulla MetrOP-tutkimus 2010–2013. Mitä tiedettiin tutkimuksen käynnistyessä keväällä 2010?*, (44–48). Helsinki: Helsingin yliopisto.
- Kuuskoski, R. (14.2.2019). Kommentti: ”Vain parasta minun lapsilleni” – koulushoppailu voi tuhota tasa-arvon mallin. Haettu osoitteesta <https://www.is.fi/kotimaa/art-2000006000697.html>
- KvantiMOTV. (2007). Mittaaminen: Muuttujien ominaisuudet. Haettu osoitteesta <https://www.fsd.tuni.fi/menetelmaopetus/mittaaminen/ominaisuudet.html>
- Lehti, H., & Laaninen, M. (2020). Perhetaustan yhteys oppimistuloksiin Suomessa PISA- ja rekisteriaineistojen valossa. <https://doi.org/10.31235/osf.io/zrhqc>
- Leivo, S. (2008). *Lukioiden valintamarkkinat opiskelijoiden opintourapolkujen ja lukioiden opintotarjousten kohtaamispaikkana: Turun päivälukiot 1998–2005*. Turku: Turun yliopisto.

- Lepistö, J. (1.10.2019). Vieraskielisten määrissä kymmenien prosenttien eroja kaupunginosien välillä – mitä huolestuttavasta kehityksestä voi seurata? *MTV*. Haettu osoitteesta <https://www.mtvuutiset.fi/artikkeli/nain-alueiden-eriytyminen-nakyy-suomessa-poismuuttoa-tietyilta-alueilta-ja-eroja-koulutusvalinnoissa-myo-uskonto-voi-aiheuttaa-yllattavan-vuokraloukun/7555042>
- Ramos Lobato, I., Bernelius, V. H., & Kosunen, S. (2018). Looking for the ordinary? Parental choice and elite school avoidance in Finland and Germany. *Nordic Journal of Studies In Educational Policy*, 4(3), (156–167). <https://doi.org/10.1080/20020317.2018.1548239>
- Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2005). *Geographic information systems and science* (2. uud. p.). Chichester: Wiley.
- Leino, K., Ahonen, A. K., Hienonen, N., Hiltunen, J., Lintuvuori, M., Lähteinen, S., Lämsä, J., Nissinen, K., Nissinen, V., Puhakka, E., Pulkkinen, J., Rautopuro, J., Sirén, M., Vainikainen, M-P. & Vettenranta, J. (2019). *PISA 18 ensituloksia*. Helsinki: Opetus- ja kulttuuriministeriö. <http://urn.fi/URN:ISBN:978-952-263-678-2>
- Marttinen, M. (14.2.2019). Suomalaisten vanhempien uusi tapa uhkaa repiä kuilun kouluihin – opettaja kertoo huolestuttavasta kehityksestä. *Iltasanomat.fi*. Haettu osoitteesta <https://www.is.fi/kotimaa/art-2000006000695.html>
- Melin, H. (26.2.2019). Suomi on edelleen luokkayhteiskunta. *Tampereen yliopiston verkkolehti Alusta!* Haettu osoitteesta <https://alusta.uta.fi/2019/02/26/suomi-on-edelleen-luokkayhteiskunta/>
- Metsämuuronen, J. (2011). *Tutkimuksen tekemisen perusteet ihmistieteissä: Tutkijalaitos* (4. korjattu laitos). Helsinki: International Methelp.
- Modin, B., Karvonen, S., Rahkonen, O., & Östberg, V. (2015). School performance, school segregation, and stress-related symptoms: Comparing Helsinki and Stockholm. *School Effectiveness and School Improvement*, 26(3), 467–486. <https://doi.org/10.1080/09243453.2014.969738>

Mukkila, S., Ilmakunnas, I., Moisio, P., & Saikkonen, P. (2017). *Perusturvan riittävyys ja köyhyys 2017*. Helsinki: Terveyden ja hyvinvoinnin laitos. <http://urn.fi/URN:ISBN:978-952-302-938-5>

Näveri, A. (2020). Etäopetus tuo mukanaan huolen eriarvoisuudesta – opetusministeri: "Koulu voi olla jollekin ainoa paikka, jossa saa lämpimän aterian". *Yle.fi*. Viitattu 4.11.2020: <https://yle.fi/uutiset/3-11261169>

Nevalainen, J. (2018). *Maahanmuuttajien asumisen ongelmat ja kehittämishaasteet*. Helsinki: Eduskunnan tarkastusvaliokunta. Haettu 13.11.2020 osoitteesta <https://www.eduskunta.fi/FI/vaski/JulkaisuMetatieto/Documents/EDK-2018-AK-169918.pdf>

Osborne, J. W. (2020). *Regression & linear modeling: Best practices and modern methods*. Thousand Oaks: SAGE Publications. <https://doi.org/10.4135/9781071802724>

Ouakrim-Soivio, N. (2013). *Toimivatko päättöarvioinnin kriteerit?: Oppilaiden saamat arvosanat ja opetushallituksen oppimistulosten seuranta-arviointi koulujen välisten osaamiserojen mittareina*. Helsinki: Opetushallitus.

Oulun kaupungin sivistys- ja kulttuurilautakunta. (2019). *Oulun kaupungin oppilaaksioton periaatteet 1.1.2020 alkaen*. Oulu: Oulun kaupunki. Haettu 13.11.2020 osoitteesta https://www.ouka.fi/documents/64277/104162/Oppilaaksioton_periaatteet_010120_alk.pdf/b9ac39a4-9159-446a-aeaa-03f5e2207c44

Oulun kaupunki. (2018). Tiernan koulu: Koulun esittely. Haettu 13.11.2020 osoitteesta <https://www.ouka.fi/fi/oulu/tiernan-koulu/koulun-esittely>

Oulun kaupunki. (2019). *Oulussa koulussa - perusopetuksen opas 2020–2021*. Oulu: Oulun kaupunki. Haettu 13.11.2020 osoitteesta https://www.ouka.fi/documents/64277/104162/Perusopetuksen_opas_2018_19_web.pdf/1ad77b1e-5d12-4356-8f06-b91243e6889a

Oulun kaupunki. (n.d.). Muuttoliike. *Ouka.fi*. Haettu 13.11.2020 osoitteesta <https://www.ouka.fi/documents/50085/15048183/2.15.pdf/8be8d484-8105-4150-bac5-7e77631a76ec>

- Pöder, K., Lauri, T., & Veski, A. (2017). Does School Admission by Zoning Affect Educational Inequality? A Study of Family Background Effect in Estonia, Finland, and Sweden. *Scandinavian Journal of Educational Research*, 61(6), 668–688. <https://doi.org/10.1080/00313831.2016.1173094>
- Rautio, S. (3.10.2019). Tutkimus Helsingistä: Perheet muuttavat pois heikko-osaisilta alueilta ennen kuin lapset menevät kouluun. Haettu osoitteesta <https://www.iltalehti.fi/kotimaa/a/1fc90321-15cc-4bb4-a3f1-45568a0fff31>
- Riihola, K. (20.9.2017). Espoon kansainvälinen koulu jäi jo pieneksi – "huippuosaajien houkuttelemiseksi koulu on tärkeä". *Länsiväylä.fi* Haettu osoitteesta <https://www.lansivayla.fi/paikalliset/1575485>
- Rossi, V. (25.1.2018). Avaaja Venla Bernelius: "Keskiluokkaisten ihmisten valinnat aiheuttavat koulujen ja alueiden eriytymisen.". *Apu.fi* Haettu osoitteesta <https://www.apu.fi/artikkelit/avaaja-venla-bernelius-keskiluokkaisten-ihmisten-valinnat-aiheuttavat-koulujen-ja>
- Saari, J. (2019). *Hyvinvointivaltio eriarvoistuneessa yhteiskunnassa: Toimi-hankkeen selvityshenkilön raportti*. Helsinki: Valtioneuvoston kanslia. <https://vnk.fi/documents/10616/5698452/Selvityshenkil%C3%B6+Juho+Saaren+raportti+30.1.2019+-+Hyvinvointivaltio+eriarvoistuneessa+yhteiskunnassa/4d051e27-3c18-7647-90c7-fd87efff4d3b/Selvityshenkil%C3%B6+Juho+Saaren+raportti+30.1.2019+-+Hyvinvointivaltio+eriarvoistuneessa+yhteiskunnassa.pdf>
- Saarimaa, T., Harjunen, O., & Kortelainen, M. (2014). Best Education Money Can Buy? Capitalization of School Quality in Finland. Helsinki: Valtion taloudellinen tutkimuskeskus. <https://doi.org/10.1093/cesifo/ifx025> Haettu osoitteesta <https://vatt.fi/documents/2956369/3012229/wp58.pdf>
- Salmela-Aro K., Chmielewski A.K. (2019) Socioeconomic Inequality and Student Outcomes in Finnish Schools. Teoksessa L. Volante, S. Schnepf, J. Jerrim, D Klinger (toim.). *Socioeconomic Inequality and Student Outcomes. Education Policy & Social Inequality*, vol 4. Singapore: Springer: https://doi.org/10.1007/978-981-13-9863-6_9 Haettu osoitteesta

https://www.researchgate.net/publication/335563675_Socioeconomic_Inequality_and_Student_Outcomes_in_Finnish_Schools

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia & Analgesia*, 126(5), (1763–1768). <https://doi.org/10.1213/ANE.0000000000002864>

Seppänen, P. (2006). *Kouluvalintapolitiikka perusopetuksessa: Suomalaiskaupunkien koulumarkkinat kansainvälisessä valossa*. (Väitöskirja). Turku: Turun yliopisto

Seppänen, P., Rinne, R., & Riipinen, P. (2012). Yläkouluvalinnat, koulujen suosio ja perheiden sosiaalinen asema: Lohkoutuuko perusopetus kaupungeissa? *Kasvatus: Suomen kasvatustieteellinen aikakauskirja*, 43(3), 226–243. <http://urn.fi/URN:NBN:fi:ELE-1564237>

Silliman, M. (2017). *Targeted Funding, Immigrant Background, and Educational Outcomes: Evidence from Helsinki's "Positive Discrimination" Policy*. Helsinki: Valtion taloudellinen tutkimuskeskus. <http://urn.fi/URN:ISBN:978-952-274-198-1>

Silvennoinen, H., Kalalahti, M., & Varjo, J. (2016). Globalisaatio, markkinaliberalismi ja koulutuspolitiikan muutos. Teoksessa H. Silvennoinen, M. Kalalahti & J. Varjo (toim.) *Koulutuksen tasa-arvon muuttuvat merkitykset* (11–34). Jyväskylä: Suomen kasvatustieteellinen seura.

Silvennoinen, H., Rinne, R., Kalalahti, M., & Varjo, J. (2015). Työväenluokan koulutusasenteet ja kouluvalinta. Teoksessa P. Seppänen, M. Kalalahti, R. Rinne & H. Simola (toim.) *Lohkoutuva peruskoulu. perheiden kouluvalinnat, yhteiskuntaluokat ja koulutuspolitiikka*. (437–468). Jyväskylä: Suomen kasvatustieteellinen seura.

Silvennoinen, H., Seppänen, P., Rinne, R., & Simola, H. (2012). Yhteiskuntaluokat ja kouluvalintapolitiikka ylikansalliselta paikalliselle tasolle ulottuvassa tarkastelussa. *Kasvatus: Suomen kasvatustieteellinen Aikakauskirja* 43(5), 502–518. <http://urn.fi/URN:NBN:fi:ELE-1581306>

- Suomen virallinen tilasto. (2016). Palkansaajien mediaaniansio 2963 euroa kuukaudessa vuonna 2015. Helsinki: Tilastokeskus. Haettu osoitteesta http://www.stat.fi/til/pra/2015/pra_2015_2016-09-29_tie_001_fi.html
- Suomen virallinen tilasto. (2018). *Vastavalmistuneiden työllisyys parani edelleen*. Helsinki: Tilastokeskus. Haettu 13.11.2020 osoitteesta https://www.stat.fi/til/sijk/2018/sijk_2018_2020-01-23_tie_001_fi.html
- Suomen virallinen tilasto. (2019). *Ammattikoulun käyneiden kokoaikatyön yleisin kuukausiansio 2 300 euroa vuonna 2018*. Helsinki: Tilastokeskus. Haettu 13.11.2020 osoitteesta https://www.stat.fi/til/pra/2018/pra_2018_2019-09-06_tie_001_fi.html
- Suomen virallinen tilasto. (2020). *Miesvaltaisilla toimialoilla suurimmat ansiot vuonna 2018*. Helsinki: Tilastokeskus. Haettu 13.11.2020 osoitteesta https://www.stat.fi/til/pra/2018/pra_2018_2020-03-20_tie_001_fi.html
- Szulkin, R., & Jonsson, J. O. (2006). Ethnic segregation and educational outcomes in Swedish comprehensive schools: A multilevel analysis. *NEP: New Economics Papers Education*. <https://www.etla.fi/edwin/publ/papers/paper10.pdf>
- Taanila, A. (2020a). *Lineaariset regressiomallit*. Haettu osoitteesta <http://taanila.fi/regressio.pdf>
- Taanila, A. (2020b). *SSPS*. Haettu osoitteesta <https://taanila.fi/spss.pdf>
- Taanila, A. (2020c). *Tilastollinen päättely*. Haettu osoitteesta <https://taanila.fi/paattely.pdf>
- Thrupp, M. (2018). Opettajat ja sosiaalinen eriarvoisuus: pitäisikö yhteiskunnallisen kontekstin merkitys unohtaa, pitäisikö sitä väheksyä vai pitäisikö se ottaa aidosti huomioon? Teoksessa R. Rinne, N. Haltia, S. Lempinen & T. Kaunisto (toim.). *Eriarvoistuva maailma – tasa-arvoistava koulu?* (47–64). Jyväskylä: Suomen kasvatustieteellinen seura.
- Tilastokeskus. (n.d.). Tilastojen ABC: 4.1 mitta-asteikot. *Stat.fi* Haettu 13.11.2020 osoitteesta https://tilastokoulu.stat.fi/verkkokoulu_v2.xql?page_type=sisalto&course_id=tkoulu_tlkt&lesson_id=4&subject_id=1

- Tilastokeskus. (2020a). Muuttoliike Oulussa 2019. *Stat.fi*. Haettu 13.11.2020 osoitteesta <https://www.stat.fi/tup/seutunet/download/businessoulu/muutot.pptx>
- Tilastokeskus. (2020b). Muuttotase kunnittain 2000–2019. *Stat.fi*. Haettu 13.11.2020 osoitteesta <https://www.stat.fi/tup/seutunet/download/businessoulu/muuttotase.pptx>
- Vaalavuo, M. (3.4.2020a). Koronavirusepidemia tuo eriarvoisuuden esiin. Blogikirjoitus. *Terveystieteiden tutkimuskeskus ja hyvinvoinnin laitos*. Haettu osoitteesta <https://blogi.thl.fi/koronavirusepidemia-tuo-eriarvoisuuden-esiin/>
- Vaalavuo, M. (3.4.2020b). Vieraskynä | koronavirusepidemia tuo eriarvoisuuden esiin. *Helsingin Sanomat*. Luettavissa myös: <https://www.hs.fi/mielipide/art-2000006461928.html>
- Vainikainen, M., Hienonen, N., Lindfors, P., Rimpelä, A., Asikainen, M., Hotulainen, R., & Hautamäki, J. (2016). Oppimistuloksia ennustavat tekijät Helsingin Metropolialueen yläkouluissa. *Kasvatus: Suomen kasvatustieteellinen aikakauskirja* 47(3), 214–229.
- Vettenranta, J., Välijärvi, J., Ahonen, A., Hautamäki, J., Hiltunen, J., Leino, K., Lähteinen, S., Nissinen, K., Nissinen, V., Puhakka, E., Rautopuro, J. & Vainikainen, M-P. (2016). *PISA 15 ensituloksia*. Helsinki: Opetus- ja kulttuuriministeriö. <http://urn.fi/URN:ISBN:978-952-263-436-8>
- Valli, R. (2015). *Johdatus tilastolliseen tutkimukseen* (2. uud. p.). Jyväskylä: PS-kustannus.
- Valtioneuvoston kanslia. (2018). *Eriarvoisuutta käsittelevän työryhmän loppuraportti*. Helsinki: Valtioneuvoston kanslia. <http://urn.fi/URN:ISBN:978-952-287-557-0>
- van Zanten, A. (2015). The determinants and dynamics of school choice: A comparative review. Teoksessa P. Seppänen, A. Carrasco, M. Kalalahti, R. Rinne & H. Simola (toim.), *Contrasting Dynamics In Education Politics of Extremes. School Choice In Chile And Finland* (3–28). Rotterdam: Sense Publishers.
- Vanas, A. (10.9.2020). Tutkittu juttu: Koulupolun alkua ei saa päästää eriytymään. *Opettaja.fi*. Haettu osoitteesta <https://www.opettaja.fi/ajassa/tutkittu-juttu-koulupolun-alkua-ei-saa-paastaa-eriytymaan/>

- Vilkama, K., & Bernelius, V. (2019). Pikkulapsiperheiden muuttoliike muovaa koulujen oppilaspohjaa Helsingissä. *Kvartti: Helsingin kaupungin tietokeskuksen neljännesvuosijulkaisu*, 2019(3), 9–14.
https://www.kvartti.fi/sites/default/files/files/issue/kvartti03_2019_verkkolehti.pdf
- Vilkama, K. & Hirvonen, J. (2018). Helsingin alueellinen eriytyminen: kaksi lähestymistapaa segregaatian seurantaan. *Kvartti: Helsingin kaupungin tietokeskuksen neljännesvuosijulkaisu*, 2018(1). Haettu osoitteesta:
<https://www.kvartti.fi/fi/artikkelit/helsingin-alueellinen-eriytyminen-kaksi-lahestymistapaa-segregaation-seurantaan>
- Vilka, H. (2014). *Tutki ja mittaa: Määrällisen tutkimuksen perusteet*. Helsinki: Tammi
- Väljärvi, J., Kupari, P., Ahonen, A. K., Arffman, I., Harju-Luukkainen, H., Leino, K., Niemivirta, M., Nissinen, K., Salmela-Aro, K., Tarnanen, M., Tuominen-Soini, H., Vettenranta, J. & Vuorinen, R. (2015). *Millä eväillä osaaminen uuteen nousuun? PISA 2012 -tutkimustuloksia*. Helsinki: Opetus- ja kulttuuriministeriö.
<http://urn.fi/URN:ISBN:978-952-263-334-7>
- Whitty, G. (2011). Markkinalähtöisyys ja sen jälkeinen aika koulutuspolitiikassa. Teoksessa R. Rinne, J. Tähtinen, A. Jauhiainen & M. Broberg (toim.). *Koulutuspolitiikan käytännöt kansallisessa ja ylikansallisessa kehyksessä* (23–43). Jyväskylä: Suomen kasvatustieteellinen seura.
- Willis, P. (1984). *Koulun penkiltä palkkatyöhön? – Miten työväenluokan nuoret saavat työväenluokan työt?* Tampere: Vastapaino.
- Yamada, I. (2017). Thiessen polygons. Teoksessa D. Richardson, N. Castree, M. F. Goodchild, A. Kobayashi, W. Liu & R. A. Marston (toim.) *International Encyclopedia of Geography: People, the Earth, Environment and Technology*. John Wiley & Sons, Ltd.
<https://doi.org/10.1002/9781118786352.wbieg0157>
- Ylpiessä, M. (07.08.2018). Näistä kunnista Ouluun muutetaan eniten – tulevaisuudessa tulomuutossa voi olla edessä iso muutos. *Kaleva*. Haettu osoitteesta
<https://www.kaleva.fi/naista-kunnista-ouluun-muutetaan-eniten-tulevaisuu/1783917>